THE

LONDON, EDINBURGH, AND DUBLIN

PHILOSOPHICAL MAGAZINE

AND

JOURNAL OF SCIENCE.

[SEVENTH SERIES.]

MARCH 1933.

XLV. *The Combustion Problem of Internal Ballistics.*—Part I.

By A. D. CROW, *O.B.E., M.A., Sc.D., F.Inst.P., and* W. E. GRIMSHAW, *O.B.E., M.A., F.Inst.P., Research Department, Woolwich* *.

INTRODUCTION.

WRITERS on internal ballistics, while differing appreciably among themselves both in basic assumptions and in methods of development, have in general been in agreement in their acceptance of the formulation of the rate of burning law for colloidal propellants in terms of the pressure only. To the importance, however, of including temperature as well as pressure explicitly as variables in a formal statement of the rate of burning law attention has been directed by Love and Pidduck † in their investigation of Lagrange's ballistic problem.

We have recently put forward formal expressions for the equation of state of a propellant gas ‡ and also for the rate of burning law for colloidal propellants § in terms of kinetic and thermodynamic quantities, and in consequence it has been thought desirable to reconsider the combustion problem in the gun *ab initio* in the light of latest developments.

* Communicated by Sir George Hadcock, F.R.S.

† Love and Pidduck, Phil. Trans. A, ccxxii. pp. 167–226 (1922).

‡ Crow and Grimshaw, Phil. Trans. A, ccxxx. 682 (1931).

§ Crow and Grimshaw, Phil. Trans. A, ccxxx. 691 (1932).

With this object in view the present investigation has been undertaken.

I. Equation Assembly.

1. *General.*—The problem to be resolved consists in the evaluation of the shot velocity, shot travel, and the gas pressure at any instant during the motion resulting from the explosive action in a gun of propellant of specified chemical composition, grain dimension, and charge weight, the internal configuration of the gun and the weight of the projectile being supposed known. Further, the treatment, to be sufficiently general, must admit provision for assessing the effects of secondary perturbing factors, the more important of which are the influence on the effective shot weight of the moving column of gas, the consequence of spin imparted to the shot by the twist of rifling, friction, wall stress, and the resistance to the motion offered by the driving band. Methods by which the necessary generalizations may be introduced will be indicated at appropriate stages in the development of the analysis.

For clarity in presentation units are based on the C. G. S. system, though any self consistent set of units could be used, as the equations ultimately to be solved present themselves in a form involving pure number only, thus enabling the ballistic and rational aspects of the problem to be considered separately.

2. *The Pressure Density Relation.*—The equation of state of a propellant gas is*

$$p = \frac{84{\cdot}80\Sigma T}{\frac{1}{\Delta'} - \eta}, \qquad \ldots\ldots (1)$$

where p is the gas pressure in kg./cm.^{2}, Σ represents Σ (gm.mols.)/gm. of the propellant gases, T is the gas temperature in °K, Δ' is the gas density in gm./cm.^{3}, and η is the co-volume in cm.^{3}/gm.

Consider a general position before complete combustion of the charge, corresponding to the instant t seconds after shot start. Defining the *size* D (cm.) of the propellant to be such that D/2 is the greatest depth burnt away below any receding surface before the whole of the grain is consumed; then, if $\phi(f)$ be the fraction of charge burnt at the instant t corresponding to a remaining thickness Df, for all normal shapes of grain

$$\phi(f) = (1-f)(1+\theta f), \qquad \ldots\ldots (2)$$

* Phil. Trans. A, ccxxx. 682, p. 72 (1931).

where $|\theta| \not> 1$, end effects being neglected. The relationship is essentially geometrical, and follows from Piobert's law of burning by parallel layers, the parameter θ taking values ranging from unity for long cords to zero for a constant burning surface, being negative and small for shapes having an increasing burning surface. Denote the charge mass in kg. by C. Then the mass of charge converted into gas at the instant t is

$$10^3 C\phi(f) \text{ gm.}$$

while the mass of solid propellant remaining at the same instant is

$$10^3 C\{1-\phi(f)\} \text{ gm.}$$

Let the density of the solid propellant be δ in gm./cm.^{3} or in kg./litre. Then the volume occupied by the unconsumed portion of the charge at the instant t is

$$10^3 C\{1-\phi(f)\}/\delta \text{ cm.}^3$$

Denoting the chamber capacity of the gun by H litres, the shot travel at the instant t by x metres, and the mean bore diameter by d cm., then the total volume behind the shot at the instant t is

$$10^3(H+Ax) \text{ cm.}^3,$$

where A is written for $\pi d^2/40$. The volume available for the propellant is therefore given by

$$10^3\left\{(H+Ax) - \frac{C\{1-\phi(f)\}}{\delta}\right\} \text{ cm.}^3,$$

and the amount of gas occupying the volume is $10^3 C\phi(f)$ gm.;

$$\therefore \quad \Delta' = \frac{C\phi(f)}{H+Ax-\frac{C}{\delta}+\frac{C\phi(f)}{\delta}} \text{ gm./cm.}^3 \text{ or kg./litre.}$$

Writing

$$E=H-C, \quad l=E/A, \quad X=(x+l)/l,$$
$$z=(1-f), \quad \text{and} \quad \nu=\theta/(1+\theta),$$

then using equation (2),

$$\Delta' = \frac{(1+\theta)C}{E} \frac{z(1-\nu z)}{X+\frac{(1+\theta)C}{E}\frac{1}{\delta}z(1-\nu z)}, \quad . \quad . \quad (3)$$

and equation (1) becomes

$$p=84{\cdot}80\Sigma T\frac{(1+\theta)C}{E}\frac{z(1-\nu z)}{X-\left(\eta-\frac{1}{\delta}\right)\frac{(1+\theta)C}{E}z(1-\nu z)}. \quad (4)$$

A digression must here be made to consider the question of energy available for imparting motion to the shot. Under ideal conditions the upper limit to the temperature would be the uncooled value T_0 calculated from the propellant composition by thermochemical methods. Owing to the imposition of wall stress there is an abstraction of energy from the propellant gases, and in consequence the temperature T_0 is never effectively operative in the gun. The work so abstracted may be taken as proportional to the product of the square of the pressure and the wall surface area exposed to the gases at any instant, the constant of proportionality depending on the build and nature of the gun. For assessing available energy values therefore we resort in the gun problem to an effective upper limit T_g to the gas temperature, less than T_0, the average numerical value of T_g throughout the motion being calculable for any particular set of circumstances. It suffices for present purposes to write $T_g=T_0/K$, where K is a constant dependent on the calibre of the gun and the maximum pressure realized. Expressing

$$\tau=T/T_g, \quad \lambda=84{\cdot}80\Sigma T_0, \quad \epsilon_3=\left(\eta-\frac{1}{\delta}\right)\frac{(1+\theta C}{E},$$

$$\Theta=\frac{\lambda}{K}\frac{(1+\theta)C}{E},$$

there results $$p=\Theta P, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where $$P=\frac{\tau z(1-\nu z)}{X-\epsilon_3 z(1-\nu z)} \quad . \quad . \quad . \quad . \quad (6)$$

an equation which is numerical.

3. *The Rate of Burning Law.*—Under closed vessel conditions it has been shown* that the combustion of a propellant is an intrinsic phenomenon, and that the rate of burning law can be formally written

$$-D\frac{df}{dt}=B\lambda\Delta', \quad . \quad . \quad . \quad . \quad . \quad (7)$$

where B and λ are constants depending on the nature of the propellant composition. The form of the equation applies

* Phil. Trans. A, ccxxx. 691 (1932).

also to the gun problem, but a modification to the *size* needs to be introduced to make up for the effect of the resistance of the driving band* at shot start and during the subsequent motion, and also to refer conditions to shot motion rather than to propellant combustion, as in the treatment that follows the origin of time is taken as the instant at which the shot starts to move. Before any movement occurs a fraction of the charge is burnt corresponding to the pressure necessary to overcome the initial resistance to shot motion. The effect of an increase in the resistance of the driving band is reflected in a corresponding increase in both the maximum pressure and the muzzle velocity, an effect which cannot be simulated analytically solely by the assignment of a shot start value for band resistance, but can be paralleled practically by firing, with no driving band, an equal charge weight of a smaller size of the same propellant composition. Equation (7) may therefore conveniently be replaced by

$$-\frac{\mathrm{D}}{k}\frac{df}{dt}=\mathrm{B}\lambda\Delta', \quad . \quad . \quad . \quad . \quad . \quad (8)$$

where the value of k can be calculated from the supposed known form of the driving band and the rifling, the equation together with $\Delta'=0$ at the start being taken to hold from the start of motion (in the ideal case of no band resistance $k=1$).

Since $f=1-z$, and all the variable involved are functions of t, equation (8) becomes

$$\frac{\mathrm{D}}{k}\frac{dz}{d\mathrm{X}}\frac{d\mathrm{X}}{dt}=\mathrm{B}\lambda\Delta'. \quad . \quad . \quad . \quad . \quad . \quad (9)$$

Let v be the shot velocity at the instant t in metres/sec. Then, writing $v=\mathrm{NV}$, where V is a pure number variable and N contains the units,

$$\mathrm{NV}=l\frac{d\mathrm{X}}{dt}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Using this relationship, and substituting the value of Δ' from equation (3) in equation (9), the rate of burning law becomes

$$\frac{d\mathrm{X}}{dz}=\frac{\mathrm{MV}\{\mathrm{X}+(\epsilon_2-\epsilon_3)z(1-\nu z\}}{z(1-\nu z)}, \quad . \quad . \quad (11)$$

where

$$\mathrm{M}=\frac{\mathrm{D}}{k}\frac{\mathrm{N}}{l}\frac{1}{\mathrm{B}\lambda}\frac{\mathrm{E}}{(1+\theta)\mathrm{C}}$$

* The driving band is an annular ring composed of relatively soft metal, attached to the shot near the base to impart spin after engraving by the rifling.

and $$\epsilon_2 = \eta(1+\theta)\frac{C}{E},$$

both being numbers. Using equation (6), this may also be written in the form

$$P\frac{dX}{dz} = MV(\tau + \epsilon_2 P), \quad . \quad . \quad . \quad . \quad (12)$$

an equation which is again purely numerical.

4. *The Equation for the Shot Motion.*—The question has to be considered of the allowance to be made for the gas motion in addition to the shot motion. It has been the general practice of writers on ballistics to assume that this can adequately be taken up by including in the effective mass, to which motion has to be imparted, a fraction of the mass of the propellant charge. The matter has been investigated by Love and Pidduck *, who found that, although the true nature of the phenomenon is extremely complicated, the conditions can be closely simulated by adding to the mass of the shot a fraction of the mass of the charge lying between the values $\frac{1}{3}$ and $\frac{1}{2}$, the gas being assumed to possess no inertia.

If, then, w be the mass of the shot in kg., and w_1 be the mass of the shot adjusted to take up for charge motion, without serious error

$$w_1 = w + \alpha C,$$

where $\frac{1}{2} > \alpha > \frac{1}{3}$.

The energy loss due to the friction between the shot and the gun-bore has also to be allowed for, and assuming that the normal friction law holds in gun conditions, the total effective mass to be moved may be written $w_1/(1-a)$, where a is the coefficient of friction, conditions now being treated as ideal.

The shot acceleration in cm./sec.^{2} is

$$10^2\frac{d^2x}{dt^2} = 10^2N^2\frac{V}{l}\frac{dV}{dX},$$

and the total thrust in dynes applied to the effective shot weight is

$$10^3 \times 980{\cdot}6 \times \frac{\pi}{4}d^2p = 10^5 \times 98{\cdot}06pA\ ;$$

$$\therefore \quad \frac{w_1}{1-a}\frac{N^2}{l}V\frac{dV}{dX} = 98{\cdot}06\ A\Theta P.$$

Since M and N are quantities at choice, subject to the

* *Loc. cit.*

relationship between them already established, a second relationship may be prescribed. Writing, therefore, since M is a number,

$$\mathrm{M} = \frac{w_1}{1-a}\frac{\mathrm{N}^2}{l}\frac{1}{98{\cdot}06}\frac{1}{\mathrm{A}}\frac{\mathrm{K}}{\lambda}\frac{\mathrm{E}}{(1+\theta)\mathrm{C}},$$

there results

$$\mathrm{MV}\frac{d\mathrm{V}}{d\mathrm{X}} = \mathrm{P}, \quad . \quad . \quad . \quad . \quad . \quad (13)$$

where

$$\mathrm{M} = \frac{98{\cdot}06}{\mathrm{K}}\left(\frac{\mathrm{AD}}{\mathrm{B}k}\right)^2\frac{1-a}{w_1}\frac{1}{\lambda\mathrm{C}}\frac{1}{1+\theta}. \quad . \quad . \quad (14)$$

It follows that

$$\mathrm{N} = \frac{98{\cdot}06\,\mathrm{AD}}{\mathrm{K}}\frac{1-a}{\mathrm{B}k}\frac{}{w_1}. \quad . \quad . \quad . \quad . \quad (15)$$

5. *Energy Considerations.*—The intrinsic energy of a gas being a function of its temperature only is ascertainable when the nature of the gas complex and the temperature are known. Denote the intrinsic energy of unit mass of the propellant gas at temperature T by $\mathrm{U_T}$. The total energy available from unit mass of gas is $\mathrm{U_{T_0}}$. By abstraction of energy due to wall stress the effective available energy per unit mass is $\mathrm{U_{T_g}}$. Of this energy the loss due to friction, twist of rifling, heat conduction, and gas motion is taken up by the employment of an effective shot mass $w_1/(1-a)$. At the instant t the gas temperature is T and the amount of propellant converted into gas is

$$10^3\mathrm{C}(1+\theta)z(1-\nu z) \text{ gm.}$$

The energy abstracted, used in the propulsion, is

$$10^3\mathrm{C}(1+\theta)z(1-\nu z)(\mathrm{U_{T_g}}-\mathrm{U_T}) \text{ gm. calories,}$$

or

$$41{\cdot}84\times10^9\mathrm{C}(1+\theta)z(1-\nu z)(\mathrm{U_{T_g}}-\mathrm{U_T}) \text{ ergs.}$$

The energy of shot motion is also equal to

$$\tfrac{1}{2}\times10^7 w_1(\mathrm{NV})^2/(1-a).$$

Over the region of temperature concerned

$$\mathrm{U_{T_g}}-\mathrm{U_T}=s(\mathrm{T_g}-\mathrm{T}),$$

where s is the mean specific heat of the gas.

It follows, using equations (14) and (15), that

$$\tau = 1-\frac{\epsilon_1\mathrm{V}^2}{z(1-\nu z)}, \quad . \quad . \quad . \quad . \quad (16)$$

where

$$\epsilon_1 = \frac{\mathrm{RM}\Sigma}{2s},$$

R being the gas constant, taken as 1·987 gm. calories.

6. *Scheme of Equations for Solution of the Rational Problem.*—For development of the analysis it is convenient to re-arrange the equations as follows :—

$$\tau = 1 - \frac{\epsilon_1 V^2}{z(1-\nu z)}, \quad . \quad . \quad . \quad . \quad . \quad (16)$$

$$\frac{dV}{dz} = \tau + \epsilon_2 P, \quad . \quad . \quad . \quad . \quad . \quad (17)$$

which follows from equations (12) and (13),

$$MV\frac{dV}{dX} = P, \quad . \quad . \quad . \quad . \quad . \quad (13)$$

and

$$P = \frac{\tau z(1-\nu z)}{X - \epsilon_3 z(1-\nu z)}, \quad . \quad . \quad . \quad . \quad . \quad (6)$$

all of which are purely numerical relationships and would be arrived at using any self consistent set of ballistic units.

II. V, X, P Functions as Weighting Numerics.

7. *General.*—The four equations under paragraph 6 involve four functions, τ, V, X, P, of one independent variable z, and by eliminating three of the functions a non-linear differential equation of the second order for V can be obtained, which does not, however, lend itself directly either to formal solution or to numerical evaluation. We proceed therefore to replace it by two differential equations of the first order, together with a third equation not involving differential coefficients. To do this τ is eliminated, using equation (16), and the three equations selected become

$$\frac{dV}{dz} = 1 - \frac{\epsilon_1 V^2}{z(1-\nu z)} + \epsilon_2 P, \quad . \quad . \quad . \quad (18)$$

$$\frac{dX}{dz} = \frac{MV\{X + (\epsilon_2 - \epsilon_3)z(1-\nu z)\}}{z(1-\nu z)}, \quad . \quad . \quad (19)$$

$$P = \frac{z(1-\nu z) - \epsilon_1 V^2}{X - \epsilon_3 z(1-\nu z)}, \quad . \quad . \quad . \quad . \quad (20)$$

8. *Numerical Evaluation of* V, X, P.—It might seem that V, X, P can be derived from the preceding equations as series in ascending powers of z. The forms, however, are unsuitable, for while V tends to a finite limit in all cases, X may reach high values. The possibility of expressing X in the form of an asymptotic expansion is discussed in Section IV. Simple direct methods, however, though unfruitful in themselves, do when generalized lead to a

systematic mode of attack for evaluating V, X, P numerically to any desired degree of accuracy. Provided h be not large and $f(z)$ be any continuous function of z,

$$f(z+h) = f(z) + \frac{h}{1!}f_1(z) + \frac{h_2}{2!}f_2(z) + \ldots, \quad . \quad (21)$$

the subscripts attaching to f denoting corresponding differentiations of the function for z. If, then, the numerical values of V, X, P as functions of z be known for the value z, the numerical values at a distance h beyond z can be computed to any required degree of accuracy provided numerical values for the differential coefficients of V, X, P for z are available at z, and the procedure can be repeated step by step.

If F, G be any two functions of z,

$$(FG)_n = F_nG + \binom{n}{1}F_{n-1}G_1 + \binom{n}{2}F_{n-2}G_2 + \ldots + \binom{n}{n}FG_n, \quad . \quad . \quad . \quad (22)$$

where FG denotes the product of the functions F and G, and $\binom{n}{r}$ the binomial coefficient

$$\frac{n!}{(n-r)!\,r!}.$$

To apply the operation of equation (22) to equations (18), (19), and (20) they are rewritten in the forms

$$z(1-\nu z)V_1 = z(1-\nu z) - \epsilon_1 V^2 + \epsilon_2 P z(1-\nu z), \quad . \quad (23)$$

$$z(1-\nu z)X_1 = MVX + M(\epsilon_2 - \epsilon_3)Vz(1-\nu z), \quad . \quad (24)$$

$$PX - \epsilon_3 z(1-\nu z)P = z(1-\nu z) - \epsilon_1 V^2. \quad . \quad . \quad . \quad . \quad (25)$$

Differentiating once for z,

$$V_2 z(1-\nu z) + \binom{1}{1}V_1(1-2\nu z) = (1-2\nu z) - \epsilon_1\{V_1V + \binom{1}{1}VV_1\} + \epsilon_2\{P_1 z(1-\nu z) + \binom{1}{1}P(1-2\nu z)\}, \quad (26)$$

$$X_2 z(1-\nu z) + \binom{1}{1}X_1(1-2\nu z) = M\{V_1X + \binom{1}{1}VX_1\} + M(\epsilon_2 - \epsilon_3)\{V_1 z(1-\nu z) + \binom{1}{1}V(1-2\nu z)\}, \quad (27)$$

$$P_1X+\binom{1}{1}PX_1-\epsilon_3\{P_1z(1-\nu z)+\binom{1}{1}P(1-2\nu z)\}$$
$$=(1-2\nu z)-\epsilon_1\{V_1V+\binom{1}{1}VV_1\}. \quad (28)$$

Differentiating twice for z,

$$V_3z(1-\nu z)+\binom{2}{1}V_2(1-2\nu z)+\binom{2}{2}V_1(-2\nu)$$
$$=-2\nu-\epsilon_1\{V_2V+\binom{2}{1}V_1V_1+\binom{2}{2}VV_2\}$$
$$+\epsilon_2\{P_2z(1-\nu z)+\binom{2}{1}P_1(1-2\nu z)+\binom{2}{2}P(-2\nu)\}, \quad (29)$$

$$X_3z(1-\nu z)+\binom{2}{1}X_2(1-2\nu z)+\binom{2}{2}X_1(-2\nu)$$
$$=M\{V_2X+\binom{2}{1}V_1X_1+\binom{2}{2}VX_2\}$$
$$+M(\epsilon_2-\epsilon_3)\{V_2z(1-\nu z)+\binom{2}{1}V_1(1-2\nu z)$$
$$+\binom{2}{2}V(-2\nu)\}, \quad (30)$$

$$P_2X+\binom{2}{1}P_1X_1+\binom{2}{2}PX_2-\epsilon_3\{P_2z(1-\nu z)$$
$$+\binom{2}{1}P_1(1-2\nu z)+\binom{2}{2}P(-2\nu)\}$$
$$=-2\nu-\epsilon_1\{V_2V+\binom{2}{1}V_1V_1+\binom{2}{2}VV_2\}. \quad (31)$$

Differentiating n times ($n>2$),

$$V_{n+1}z(1-\nu z)+\binom{n}{1}V_n(1-2\nu z)+\binom{n}{2}V_{n-1}(-2\nu)$$
$$=-\epsilon_1\{V_nV+\binom{n}{1}V_{n-1}V_1+\ldots\ldots+\binom{n}{n}VV_n\}$$
$$+\epsilon_2\{P_nz(1-\nu z)+\binom{n}{1}P_{n-1}(1-2\nu z)$$
$$+\binom{n}{2}P_{n-2}(-2\nu)\}, \quad (32)$$

$$X_{n+1}z(1-\nu z)+\binom{n}{1}X_n(1-2\nu z)+\binom{n}{2}X_{n-1}(-2\nu)$$
$$=M\{V_nX+\binom{n}{1}V_{n-1}X_1+\ldots\ldots+\binom{n}{n}VX_n\}$$
$$+M(\epsilon_2-\epsilon_3)\{V_nz(1-\nu z)+\binom{n}{1}V_{n-1}(1-2\nu z)$$
$$+\binom{n}{2}V_{n-2}(-2\nu)\}, \quad . \quad (33)$$

$$P_nX+\binom{n}{1}P_{n-1}X_1+\ldots\ldots+\binom{n}{n}PX_n-\epsilon_3\{P_nz(1-\nu z)$$
$$+\binom{n}{1}P_{n-1}(1-2\nu z)+\binom{n}{2}P_{n-2}(-2\nu)\}$$
$$=-\epsilon_1\{V_nV+\binom{n}{1}V_{n-1}V_1+\ldots\ldots+\binom{n}{n}VV_n\}. \quad (34)$$

The method as developed applies for finite values of z. At the start for $z=0$ a modified procedure must be adopted, two methods being available: first, the terms containing the ϵ parameters are relatively not large, and may be neglected, provided z be sufficiently small.

Equations (18), (19), and (20) then become

$$\frac{dV}{dz}=1,$$
$$\frac{dX}{dz}=\frac{MVX}{z(1-\nu z)},$$
$$P=\frac{z(1-\nu z)}{X},$$

the starting conditions being $z=0$, $V=0$, $X=1$, $P=0$, so that approximate solutions are:

$$V=z,$$
$$X=(1-\nu z)^{-\frac{M}{\nu}},$$
$$P=z(1-\nu z)^{\frac{M}{\nu}+1}.$$

The method, however, is unsatisfactory in that no indication is afforded of the magnitude of quantities omitted.

To overcome the difficulty and to arrange for expressions in a manner to keep control of the orders of magnitude neglected V, X, P are found as series in ascending powers of z, with the restriction that z is not to be large. Evaluating,

then, the differential coefficients for $z=0$, from equations (26), (27), and (28),

$$V_1=1,$$
$$X_1=M,$$
$$P_1=1.$$

Equations (29), (30), and (31) give

$$V_2=-(\epsilon_1-\epsilon_2),$$
$$X_2=M(M+\nu-\frac{1}{2}\epsilon_1+\frac{3}{2}\epsilon_2-\epsilon_3),$$
$$P_2=-2(M+\nu+\epsilon_1-\epsilon_3).$$

The method of arriving at numerical values of higher derivatives of the functions follows from the use of special cases of equations (32), (33), and (34), and the procedure for computing numerical values of V, X, P for any value of z, where z is not large, is established. Numerical values being assigned to M, ν, ϵ_1, ϵ_2, ϵ_3, coefficients of terms up to the second powers of z are immediately obtained, values and coefficients for high powers being then derived from the established linkages ; the procedure being discontinued as soon as inspection shows that to the order of accuracy required further coefficients can be neglected. As z varies from 0 to 1, *i. e.*, from the start to the end of combustion, V and X progressively increase, while P in the generality of cases increases to a maximum and then decreases, the occurrence of this phenomenon depending on the relative rates of increase of the terms in the numerator and denominator in the expression for P. The value of τ immediately follows from equation (16) when V is known.

III. Solution by (q, r) Function Loadings of ϵ Parameter Powers and Products.

9. *Preliminary.*—The scheme of equations in paragraph 6 includes a basic parameter M and three subsidiary parameters ϵ_1, ϵ_2, ϵ_3, the terms attaching to which never attain very high values compared with other magnitudes involved. The ballistic functions may therefore be expressed in terms of series in ascending powers of the ϵ parameters, the coefficients of which are functions of the independent variable z, and as such act as weighting factors in assessing the effects due to the respective subsidiary parameters. It is seldom necessary to derive a solution beyond the first power of ϵ_1, ϵ_2, ϵ_3. For completeness, however, the formal solution is

extended to second powers and products, terms beyond these being of the third order of small quantities, and negligible. The analysis in effect deals with perturbations imposed on a primitive solution of the problem.

10. *Rearrangement of the Primary Equations.*

Let

$$V=[V_0]+\epsilon_1[V_1]+\epsilon_2[V_2]+\epsilon_3[V_3]+\epsilon_1^2[V_{11}]+\epsilon_2^2[V_{22}]+\epsilon_3^2[V_{33}]+\epsilon_1\epsilon_2[V_{12}]+\epsilon_2\epsilon_3[V_{23}]+\epsilon_3\epsilon_1[V_{31}], \quad (35)$$

$$\log X=\log[X_0]+\epsilon_1[X_1]+\epsilon_2[X_2]+\epsilon_3[X_3]+\epsilon_1^2[X_{11}]+\epsilon_2^2[X_{22}[+\epsilon_3^2[X_{33}]+\epsilon_1\epsilon_2[X_{12}]+\epsilon_2\epsilon_3[X_{23}]+\epsilon_3\epsilon_1[X_{31}], \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (36)$$

$$\log P=\log P_0+\epsilon_1P_1+\epsilon_2P_2+\epsilon_3P_3+\epsilon_1^2P_{11}+\epsilon_2^2P_{22}+\epsilon_3^2P_{33}+\epsilon_1\epsilon_2P_{12}+\epsilon_2\epsilon_3P_{23}+\epsilon_3\epsilon_1P_{31}, \quad . \quad . \quad . \quad (37)$$

where the suffixed V, X, P functions contain M, ν, z only and the enclosure of any quantity within square brackets denotes the integration from 0 to z of the quantity concerned *.

Using equation (37) and equating coefficients of powers and products of the parameters, equation (18) yields:

$$V_0=1, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (38)(\text{i.})$$

$$V_1=-\frac{[V_0]^2}{z(1-\nu z)}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (38)(\text{ii.})$$

$$V_2=P_0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (38)(\text{iii.})$$

$$V_3=0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (38)(\text{iv.})$$

$$V_{11}=-\frac{2[V_0][V_1]}{z(1-\nu z)}, \quad . \quad . \quad . \quad . \quad (38)(\text{v.})$$

$$V_{22}=P_0P_2, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (38)(\text{vi.})$$

$$V_{33}=0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (38)(\text{vii.})$$

$$V_{12}=-\frac{2[V_0][V_2]}{z(1-\nu z)}+P_0P_1, \quad . \quad (38)(\text{viii.})$$

$$V_{23}=P_0P_3, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (38)(\text{ix.})$$

$$V_{31}=-\frac{2[V_0][V_3]}{z(1-\nu z)}, \quad . \quad . \quad . \quad (38)(\text{x.})$$

* No confusion need arise between the present notation and that of Section II.

Similarly using equations (19) and (36):

$$X_0 = \frac{M[X_0][V_0]}{z(1-\nu z)}, \quad \ldots \quad (39)(\text{i.})$$

$$X_1 = \frac{M[V_1]}{z(1-\nu z)}, \quad \ldots \quad (39)(\text{ii.})$$

$$X_2 = \frac{M[V_2]}{z(1-\nu z)} + \frac{M[V_0]}{[X_0]}, \quad \ldots \quad (39)(\text{iii.})$$

$$X_3 = \frac{M[V_3]}{z(1-\nu z)} - \frac{M[V_0]}{[X_0]}, \quad \ldots \quad (39)(\text{iv.})$$

$$X_{11} = \frac{M[V_{11}]}{z(1-\nu z)}, \quad \ldots \quad (39)(\text{v.})$$

$$X_{22} = \frac{M[V_{22}]}{z(1-\nu z)} + \frac{M\{[V_2]-[V_0][X_2]\}}{[X_0]}, \quad \ldots \quad (39)(\text{vi.})$$

$$X_{33} = \frac{M[V_{33}]}{z(1-\nu z)} - \frac{M\{[V_3]-[V_0][X_3]\}}{[X_0]}, \quad \ldots \quad (39)(\text{vii.})$$

$$X_{12} = \frac{M[V_{12}]}{z(1-\nu z)} + \frac{M\{[V_1]-[V_0][X_1]\}}{[X_0]}, \quad \ldots \quad (39)(\text{viii.})$$

$$X_{23} = \frac{M[V_{23}]}{z(1-\nu z)} - \frac{M\{[V_2]-[V_0][X_2]-[V_3]+[V_0][X_3]\}}{[X_0]}, \quad \ldots \quad (39)(\text{ix.})$$

$$X_{31} = \frac{M[V_{31}]}{z(1-\nu z)} - \frac{M\{[V_1]-[V_0][X_1]\}}{[X_0]}. \quad \ldots \quad (39)(\text{x.})$$

Finally, from equations (20) and (37):

$$P_0 = \frac{z(1-\nu z)}{[X_0]}, \quad \ldots \quad (40)(\text{i.})$$

$$P_1 = -\frac{[V_0]^2}{[X_0]P_0} - [X_1], \quad \ldots \quad (40)(\text{ii.})$$

$$P_2 = -[X_2], \quad \ldots \quad (40)(\text{iii.})$$

$$P_3 = -[X_3] + \frac{z(1-\nu z)}{[X_0]}, \quad \ldots \quad (40)(\text{iv.})$$

$$P_{11} = -\frac{2[V_0][V_1]}{[X_0]P_0} - \frac{[X_1]^2}{2} - [X_{11}] - [X_1]P_1 - \frac{P_1^2}{2}, \quad \ldots \quad (40)(\text{v.})$$

$$P_{22} = -\frac{[X_2]^2}{2} - [X_{22}] - [X_2]P_2 - \frac{P_2^2}{2}, \quad \ldots \quad (40)(\text{vi.})$$

$$P_{33}=-\frac{[X_3]^2}{2}-[X_{33}]-[X_3]P_3-\frac{P_3^2}{2}+\frac{P_3z(1-\nu z)}{[X_0]}, \quad . \ . \ . \quad (40)(\text{vii.})$$

$$P_{12}=-\frac{2[V_0][V_2]}{[X_0]P_0}-[X_1][X_2]-[X_{12}]-[X_1]P_2 - [X_2]P_1-P_1P_2, \quad . \ . \ . \quad (40)(\text{viii.})$$

$$P_{23}=-[X_2][X_3]-[X_{23}]-[X_2]P_3-[X_3]P_2-P_2P_3 + \frac{P_2z(1-\nu z)}{[X_0]}, \quad . \ . \quad (40)(\text{ix.})$$

$$P_{31}=-\frac{2[V_0][V_3]}{[X_0]P_0}-[X_3][X_1]-[X_{31}]-[X_3]P_1-[X_1]P_3 - P_3P_1+\frac{P_1z(1-\nu z)}{[X_0]}. \quad . \ . \quad (40)(\text{x.})$$

11. *Convergency Considerations.*—In what follows the expression $z^q(1-\nu z)^r$ will be denoted by (q, r), and the integrated expression between 0 and z by $[(q, r)]$, according to the convention of paragraph (10).

The following relationships hold:

$$(q, r)(s, t)=(q+s, r+t), \quad . \ . \ . \ . \ . \ . \ . \ . \ . \quad (41)$$

$$[(q, r)]=\frac{1}{q+1}(q+1, r)+\frac{r\nu}{q+1}[(q+1, r-1)]. \quad (42)$$

From equation (42) there follows:

$$[(q, r)]=\binom{q, r}{1}(q+1, r)+\binom{q, r}{2}(q+2, r-1)+\ldots + \binom{q, r}{n}(q+n, r-n+1)+\ldots, \quad (43)$$

where

$$\binom{q, r}{1}=\frac{1}{q+1},$$

the general coefficient being given by

$$\binom{q, r}{n}=\frac{r\nu(r-1)\nu(r-2)\nu\ldots(r-n+2)\nu}{(q+1)(q+2)\ldots(q+n)}.$$

Resort to series of the type given in equation (43) is compelled by reason of the repeated occurrence of the expression $(-1, -1)$ in quantities requiring to be integrated, and the convergency conditions of such series must in consequence be examined. From the left-hand side of equation (43), since a summation of positive quantities only is involved, $[(q, r)]$ increases with z attaining a maximum

value for $z=1$. The right-hand side of equation (43) may be written

$$(q+1, r)\left\{\binom{q, r}{1}+\binom{q, r}{2}(1, -1)+\binom{q, r}{3}(2, -2)+ \cdots + \binom{q, r}{n+1}(n, -n)+ \cdots \right\},$$

and

$$\lim_{n\to\infty} \frac{\binom{q, r}{n+1}(n, -n)}{\binom{q, r}{n}(n-1, -n+1)} = \lim_{n\to\infty} \frac{(r-n+1)\nu}{q+n+1}(1, -1).$$

Now $|\nu(1, -1)| < 1$ for $z < 1$, and also for $z=1$ for all values of $\nu < \frac{1}{2}$, the limiting value of ν that can occur. It follows therefore that for $\nu<\frac{1}{2}$ the series is convergent for all values of z from 0 to 1. In the special case of $\nu=\frac{1}{2}$, $\nu(1, -1)=1$, and the expression within brackets to be summed is

$$\frac{1}{q+1}+\frac{r}{(q+1)(q+2)}+\frac{r(r-1)}{(q+1)(q+2)(q+3)}+ \cdots + \frac{r(r-1)(r-2) \ldots (r-n)}{(q+1)(q+2)(q+3) \ldots (q+n+2)}+ \ldots . \quad (44)$$

The quantity q is always positive and integral, while r may take positive or negative values, it being found in the latter case that $|r| \not> (q+1)$. The terms of (44) with r negative are alternately positive and negative from the start, and the convergency at slowest is comparable with that of the normal expansion for $\log_e 2$, as is illustrated subsequently by consideration of the case $q=1$, $r=-1$.

In the general case of r positive let s be the integer next below r, so that up to and including the $(s+2)$th term all the terms are positive and rapidly decreasing in value, the signs thenceforward being alternately positive and negative. Considering then the remainder terms, starting with the one immediately prior to the occurrence of the first negative sign,

$$\frac{r(r-1)(r-2) \ldots (r-s)}{(q+1)(q+2)(q+3) \ldots (q+s+2)} + \frac{r(r-1)(r-2) \ldots (r-s-1)}{(q+1)(q+2)(q+3) \ldots (q+s+3)}+ \cdots + \frac{r(r-1)(r-2) \ldots (r-s-t)}{(q+1)(q+2)(q+3) \ldots (q+s+t+2)}+ \ldots,$$

the expression can be written in the form

$$\frac{r(r-1)(r-2)\ \ldots\ (r-s)}{(q+1)(q+2)(q+3)\ \ldots\ (q+s+2)}$$
$$\left\{\left(1+\frac{r-s-1}{q+s+3}\right)+\frac{(r-s-1)(r-s-2)}{(q+s+3)(q+s+4)}\right.$$
$$\left.\left(1+\frac{r-s-3}{q+s+5}\right)+\ \ldots\ \right\},$$

which is positive, and also in the form

$$\frac{r(r-1)(r-2)\ \ldots\ (r-s)}{(q+1)(q+2)(q+3)\ \ldots\ (q+s+2)}$$
$$\left\{1+\frac{r-s-1}{q+s+3}\left(1+\frac{r-s-2}{q+s+4}\right)\right.$$
$$+\frac{(r-s-1)(r-s-2)(r-s-3)}{(q+s+3)(q+s+4)(q+s+5)}$$
$$\left.\left(1+\frac{r-s-4}{q+s+6}\right)+\ \ldots\ \right\},$$

showing that the series within the brackets is less than unity, each term afetr the first being negative.

Finally, in the special case of $\nu=\frac{1}{2}$, $q=1$, $r=-1$, for $z=1$ series (44) becomes

$$\frac{1}{1.2}-\frac{1}{2.3}+\frac{1}{3.4}-\frac{1}{4.5}+\ \ldots,\ .\ .\ .\qquad(45)$$

the convergency of which is comparable with $\Sigma(1/n^2)$, the sum being $2\log_e 2-1$. In every case therefore series of the type in equation (43) occurring in the subsequent analysis are absolutely and uniformly convergent; and as the ϵ coefficients are obtained by the integration of sums of powers and products of such series it follows that they will all be finite quantities. In practice convergency is rapid, and the summed quantities are not large by reason of the repeated integrations.

12. *Forms of the ϵ Coefficients.*—We may now proceed to the solution of the subsidiary equations under (38), (39), and (40). The order of evaluation consists in first deriving as many solutions as possible of the (38) equations, proceeding thence to the (39) equations, onwards to the (40) equations, reverting to the (38) equations, and so on. The explicit forms for the ϵ coefficients for V and X are given below, the form of the P functions following directly when these are known from equation (20).

[V] *Functions*:

$$[V_0] = (1, 0),$$

$$[V_1] = -[(1, -1)],$$

$$[V_2] = \left[\left(1, \frac{M}{\nu} + 1\right)\right],$$

$$[V_3] = 0,$$

$$[V_{11}] = 2[\,[(1, -1)]\,(0, -1)],$$

$$[V_{22}] = -M\left[\,\left[\left(1, \frac{M}{\nu}\right)\right]\left(1, \frac{M}{\nu} + 1\right)\right]$$

$$-M\left[\,\left[\,\left[\left(1, \frac{M}{\nu} + 1\right)\right](-1, -1)\right]\left(1, \frac{M}{\nu} + 1\right)\right],$$

$$[V_{33}] = 0,$$

$$[V_{12}] = -\left[\left(2, \frac{M}{\nu}\right)\right] - 2\left[\,\left[\left(1, \frac{M}{\nu} + 1\right)\right](0, -1)\right]$$

$$+M\left[\,[\,[(1, -1)]\,(-1, -1)]\left(1, \frac{M}{\nu} + 1\right)\right],$$

$$[V_{23}] = \left[\left(2, \frac{2M}{\nu} + 2\right)\right] + M\left[\,\left[\left(1, \frac{M}{\nu}\right)\right]\left(1, \frac{M}{\nu} + 1\right)\right],$$

$$[V_{31}] = 0.$$

[X] *Functions*:

$$[X_0] = \left(0, -\frac{M}{\nu}\right),$$

$$[X_1] = -M\,[\,[(1, -1)]\,(-1, -1)],$$

$$[X_2] = M\left[\left(1, \frac{M}{\nu}\right)\right] + M\left[\,\left[\left(1, \frac{M}{\nu} + 1\right)\right](-1, -1)\right],$$

$$[X_3] = -M\left[\left(1, \frac{M}{\nu}\right)\right],$$

$$[X_{11}] = 2M\,[\,[\,[(1, -1)]\,(0, -1)]\,(-1, -1)],$$

$$[X_{22}] = M\left[\,\left[\left(1, \frac{M}{\nu} + 1\right)\right]\left(0, \frac{M}{\nu}\right)\right]$$

$$-M^2\left[\,\left[\left(1, \frac{M}{\nu}\right)\right]\left(1, \frac{M}{\nu}\right)\right]$$

$$-M^2\Big[\Big[\Big[\Big(1,\frac{M}{\nu}\Big)\Big]\Big(1,\frac{M}{\nu}+1\Big)\Big](-1,-1)\Big]$$

$$-M^2\Big[\Big[\Big[\Big(1,\frac{M}{\nu}+1\Big)\Big](-1,-1)\Big]\Big(1,\frac{M}{\nu}\Big)\Big]$$

$$-M^2\Big[\Big[\Big[\Big[\Big(1,\frac{M}{\nu}+1\Big)\Big](-1,-1)\Big]\Big(1,\frac{M}{\nu}+1\Big)\Big](-1,-1)\Big],$$

$$[X_{33}]=-M^2\Big[\Big[\Big(1,\frac{M}{\nu}\Big)\Big]\Big(1,\frac{M}{\nu}\Big)\Big],$$

$$[X_{12}]=-M\Big[[(1,-1)]\Big(0,\frac{M}{\nu}\Big)\Big]$$

$$-M\Big[\Big[\Big(2,\frac{M}{\nu}\Big)\Big](-1,-1)\Big]$$

$$-2M\Big[\Big[\Big[\Big(1,\frac{M}{\nu}+1\Big)\Big](0,-1)\Big](-1,-1)\Big]$$

$$+M^2\Big[[[(1,-1)](-1,-1)]\Big(1,\frac{M}{\nu}\Big)\Big]$$

$$+M^2\Big[\Big[[[(1,-1)](-1,-1)]\Big(1,\frac{M}{\nu}+1\Big)\Big](-1,-1)\Big],$$

$$[X_{23}]=-M\Big[\Big[\Big(1,\frac{M}{\nu}+1\Big)\Big]\Big(0,\frac{M}{\nu}\Big)\Big]$$

$$+M\Big[\Big[\Big(2,\frac{2M}{\nu}+2\Big)\Big](-1,-1)\Big]$$

$$+2M^2\Big[\Big[\Big(1,\frac{M}{\nu}\Big)\Big]\Big(1,\frac{M}{\nu}\Big)\Big]$$

$$+M^2\Big[\Big[\Big[\Big(1,\frac{M}{\nu}\Big)\Big]\Big(1,\frac{M}{\nu}+1\Big)\Big](-1,-1)\Big]$$

$$+M^2\Big[\Big[\Big[\Big(1,\frac{M}{\nu}+1\Big)\Big](-1,-1)\Big]\Big(1,\frac{M}{\nu}\Big)\Big],$$

$$[X_{31}] = M\left[[(1,-1)]\left(0, \frac{M}{\nu}\right)\right]$$

$$-M^2\left[[[(1,-1)](-1,-1)]\left(1, \frac{M}{\nu}\right)\right].$$

13. *Evaluation of the Coefficients.*—Inspection of the coefficients tabulated above shows that quantities only are involved of the form (q, r), $[(q, r)]$ together with products and integrated functions. The repeated integrations can be most compactly achieved by a method involving repeated differentiation, simplification, and establishment of recurrence formulæ amongst coefficients. Consider first the function (q, r) and denote n differentiations of this for z by $(q, r)_n$.

Then

$$(q, r)_1 = q(q-1, r) - \nu r(q, r-1).$$

The operation

$$(\ ,\ )_1 \equiv 0_{(1)} - \nu 0_{(2)},$$

where the operator $0_{(1)}$ reduces the first numeric *inside* (q, r) by unity and multiplies by the first numeric, commuting with the second numeric, while the operator $0_{(2)}$ reduces the second numeric *inside* (q, r) by unity and multiplies by the second numeric, commuting with the first numeric.

$$\therefore (\ ,\ )_n \equiv \{0_{(1)} - \nu 0_{(2)}\}^n (q, r)$$

$$\equiv \{(0_{(1)})^n - \binom{n}{1}\nu (0_{(1)})^{n-1} 0_{(2)}$$

$$+ \binom{n}{2}\nu^2 (0_{(1)})^{n-2} (0_{(2)})^2 + \dots\}(q, r),$$

the indices of the 0 terms denoting repeated operations and $\binom{n}{s}$ the ordinary binominal coefficient.

$$\therefore (q, r)_n = q(q-1)(q-2)\dots(q-n+1)(q-n, r)$$

$$-\binom{n}{1}\nu q(q-1)(q-2)\dots(q-n+2)r(q-n+1, r-1)$$

$$+\binom{n}{2}\nu^2 q(q-1)(q-2)\dots(q-n+3)r(r-1)(q-n+2, r-2)$$

$$-\ \dots\dots\dots\dots\dots\dots\dots\dots$$

As an illustration of general procedure in performing repeated integrations consider the expression

$$[[[[(q_1, r_1)](q_2, r_2)](q_3, r_3)](q_4, r_4)].$$

Inspection shows that the first term is of the form

$$(q_1+q_2+q_3+q_4,\ r_1+r_2+r_3+r_4),$$

and the expression may be written

$$\sum_1 C_n(q_1+q_2+q_3+q_4+n+3,\ r_1+r_2+r_3+r_4-n+1),$$

where the coefficients C_n remain to be determined. By repeated differentiation and simplification the expression reduces to

$$1 = \sum_1^\infty C_n \left\{ \left\{ \begin{matrix} 0 \\ n \end{matrix} \right\} (n-1,\ -n+1) - \nu \left\{ \begin{matrix} 1 \\ n \end{matrix} \right\} (n,\ -n) \right.$$
$$+ \nu^2 \left\{ \begin{matrix} 2 \\ n \end{matrix} \right\} (n+1,\ -n-1) - \nu^3 \left\{ \begin{matrix} 3 \\ n \end{matrix} \right\} (n+2,\ -n-2)$$
$$\left. + \nu^4 \left\{ \begin{matrix} 4 \\ n \end{matrix} \right\} (n+3,\ -n-3) \right\},$$

where

$$\left\{ \begin{matrix} 0 \\ n \end{matrix} \right\} = (q_4+n+3)(q_3+n+2)(q_2+n+1)(q_1+n),$$

$$\left\{ \begin{matrix} 1 \\ n \end{matrix} \right\} = (q_4+n+3)(q_3+n+2)(q_2+n+1)(r_1-n+1)$$
$$+ (q_4+n+3)(q_3+n+2)(q_1+n+1)(r_2-n+1)$$
$$+ (q_4+n+3)(q_2+n+2)(q_1+n+1)(r_3-n+1)$$
$$+ (q_3+n+3)(q_2+n+2)(q_1+n+1)(r_4-n+1),$$

$$\left\{ \begin{matrix} 2 \\ n \end{matrix} \right\} = (q_4+n+3)(q_3+n+2)(r_2-n+1)(r_1-n)$$
$$+ (q_4+n+3)(q_2+n+2)(r_3-n+1)(r_1-n)$$
$$+ (q_4+n+3)(q_1+n+2)(r_3-n+1)(r_2-n)$$
$$+ (q_3+n+3)(q_2+n+2)(r_4-n+1)(r_1-n)$$
$$+ (q_3+n+3)(q_1+n+2)(r_4-n+1)(r_2-n)$$
$$+ (q_2+n+3)(q_1+n+2)(r_4-n+1)(r_3-n),$$

$$\left\{\begin{matrix}3\\n\end{matrix}\right\} = (\bar{q}_4+n+3)(\bar{r}_3-n+1)(\bar{r}_2-n)(r_1-n-1)$$
$$+(\bar{q}_3+n+3)(r_4-n+1)(\bar{r}_2-n)(r_1-n-1)$$
$$+(\bar{q}_2+n+3)(\bar{r}_4-n+1)(\bar{r}_3-n)(r_1-n-1)$$
$$+(q_1+n+3)(\bar{r}_4-n+1)(\bar{r}_3-n)(\bar{r}_2-n-1),$$

$$\left\{\begin{matrix}4\\n\end{matrix}\right\} = (\bar{r}_4-n+1)(\bar{r}_3-n)(\bar{r}_2-n-1)(r_1-n-2),$$

the symbols $\bar{q}_s$ and $\bar{r}_s$ representing respectively sums of q and r from 1 to s.

The relationships giving the respective coefficients are then as follows :—

$$1 = \left\{\begin{matrix}0\\1\end{matrix}\right\} C_1,$$

$$0 = \left\{\begin{matrix}0\\2\end{matrix}\right\} C_2 - \nu \left\{\begin{matrix}1\\1\end{matrix}\right\} C_1,$$

$$0 = \left\{\begin{matrix}0\\3\end{matrix}\right\} C_3 - \nu \left\{\begin{matrix}1\\2\end{matrix}\right\} C_2 + \nu^2 \left\{\begin{matrix}2\\1\end{matrix}\right\} C_1,$$

$$0 = \left\{\begin{matrix}0\\4\end{matrix}\right\} C_4 - \nu \left\{\begin{matrix}1\\3\end{matrix}\right\} C_3 + \nu^2 \left\{\begin{matrix}2\\2\end{matrix}\right\} C_2 - \nu^3 \left\{\begin{matrix}3\\1\end{matrix}\right\} C_1,$$

$$0 = \left\{\begin{matrix}0\\5\end{matrix}\right\} C_5 - \nu \left\{\begin{matrix}1\\4\end{matrix}\right\} C_4 + \nu^2 \left\{\begin{matrix}2\\3\end{matrix}\right\} C_3 - \nu^3 \left\{\begin{matrix}3\\2\end{matrix}\right\} C_2 + \nu^4 \left\{\begin{matrix}4\\1\end{matrix}\right\} C_1,$$

and, generally,

$$0 = \left\{\begin{matrix}0\\n\end{matrix}\right\} C_n - \nu \left\{\begin{matrix}1\\n-1\end{matrix}\right\} C_{n-1}$$
$$+\nu^2 \left\{\begin{matrix}2\\n-2\end{matrix}\right\} C_{n-2} - \nu^3 \left\{\begin{matrix}3\\n-3\end{matrix}\right\} C_{n-3} + \nu^4 \left\{\begin{matrix}4\\n-4\end{matrix}\right\} C_{n-4}.$$

The forms obtained render direct the numerical evaluation of the integrals.

Explicitly expressed,

$$[[[[(q_1, r_1)](q_2, r_2)](q_3, r_3)](q_4, r_4)]$$
$$= \sum_1^\infty \frac{\nu^{n-1}\Delta_n}{\left\{\begin{matrix}0\\1\end{matrix}\right\}\left\{\begin{matrix}0\\2\end{matrix}\right\}\left\{\begin{matrix}0\\3\end{matrix}\right\}\cdots\left\{\begin{matrix}0\\n\end{matrix}\right\}} (\bar{q}_4+n+3, \bar{r}_4-n+1), \quad (46)$$

where $\Delta_1 = 1,$

$$\Delta_2 = \left\{\begin{matrix}1\\1\end{matrix}\right\},$$

and for $n>2$

$$\Delta_n = \begin{vmatrix}
\left\{\begin{matrix}1\\n-1\end{matrix}\right\}, & -\left\{\begin{matrix}2\\n-2\end{matrix}\right\}, & \left\{\begin{matrix}3\\n-3\end{matrix}\right\}, & -\left\{\begin{matrix}4\\n-4\end{matrix}\right\}, & 0, & \cdots & 0, \\
-\left\{\begin{matrix}0\\n-1\end{matrix}\right\}, & \left\{\begin{matrix}1\\n-2\end{matrix}\right\}, & -\left\{\begin{matrix}2\\n-3\end{matrix}\right\}, & \left\{\begin{matrix}3\\n-4\end{matrix}\right\}, & -\left\{\begin{matrix}4\\n-5\end{matrix}\right\}, & \cdots & \\
0, & -\left\{\begin{matrix}0\\n-2\end{matrix}\right\}, & \left\{\begin{matrix}1\\n-3\end{matrix}\right\}, & -\left\{\begin{matrix}2\\n-4\end{matrix}\right\}, & \left\{\begin{matrix}3\\n-5\end{matrix}\right\}, & \cdots & \\
0, & 0, & -\left\{\begin{matrix}0\\n-3\end{matrix}\right\}, & \left\{\begin{matrix}1\\n-4\end{matrix}\right\}, & -\left\{\begin{matrix}2\\n-5\end{matrix}\right\}, & \cdots & \\
0, & 0, & 0, & -\left\{\begin{matrix}0\\n-4\end{matrix}\right\}, & \left\{\begin{matrix}1\\n-5\end{matrix}\right\}, & \cdots & \\
0, & 0, & 0, & 0, & -\left\{\begin{matrix}0\\n-5\end{matrix}\right\}, & \cdots & \\
0, & 0, & 0, & 0, & 0, & \cdots & \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0, \\
\cdots & \cdots & \cdots & \cdots & \cdots & -\left\{\begin{matrix}0\\5\end{matrix}\right\}, & \left\{\begin{matrix}1\\4\end{matrix}\right\}, \\
\cdots & \cdots & \cdots & \cdots & \cdots & 0, & -\left\{\begin{matrix}0\\4\end{matrix}\right\}, \\
\cdots & \cdots & \cdots & \cdots & \cdots & 0, & 0, \\
0, & \cdots & \cdots & \cdots & \cdots & 0, & 0,
\end{vmatrix}$$

$$\begin{matrix}
-\left\{\begin{matrix}2\\3\end{matrix}\right\}, & \left\{\begin{matrix}3\\2\end{matrix}\right\}, & -\left\{\begin{matrix}4\\1\end{matrix}\right\}, \\
\left\{\begin{matrix}1\\3\end{matrix}\right\}, & -\left\{\begin{matrix}2\\2\end{matrix}\right\}, & \left\{\begin{matrix}3\\1\end{matrix}\right\}, \\
-\left\{\begin{matrix}0\\3\end{matrix}\right\}, & \left\{\begin{matrix}1\\2\end{matrix}\right\}, & -\left\{\begin{matrix}2\\1\end{matrix}\right\}, \\
0, & -\left\{\begin{matrix}0\\2\end{matrix}\right\}, & \left\{\begin{matrix}1\\1\end{matrix}\right\}.
\end{matrix}$$

The series in equation (46) does not necessarily end with the first term when $\nu=0$, for whilst the q's are normally positive integers, the r's may comprise values of the form $\frac{a\mathrm{M}}{\nu}$, where a is a positive integer, leading, for example, to

$$\lim_{\nu=0}\left(q, \frac{a\mathrm{M}}{\nu}\right) = z^q e^{-a\mathrm{M}z},$$

and $\lim_{\nu=0} \nu^{n-1}\Delta_{n-1}$ finite for all n, the repeated factorials in the denominators of the terms on the right-hand side ensuring convergency. The case of $[\mathrm{X}_1]$ for $\nu=\frac{1}{2}$, $z=1$ may be considered as exemplifying numerical values and orders of magnitude.

By equation (43)

$$[(1,-1)] = \frac{1}{2}(2,-1) - \frac{\nu}{2.3}(3,-2) + \frac{\nu.2\nu}{2.3.4}(4,-3) - \frac{\nu.2\nu.3\nu}{2.3.4.5}(5,-4) + \ldots,$$

so that

$$[(1,-1)](-1,-1) = \frac{1}{2}(1,-2) - \frac{\nu}{2.3}(2,-3) + \frac{\nu.2\nu}{2.3.4}(3,-4) - \frac{\nu.2\nu.3\nu}{2.3.4.5}(4,-5) + \ldots,$$

and

$$\begin{aligned}[[(1,-1)](-1,-1)] &= \frac{1}{2}\left\{\frac{1}{2}(2,-2) - \frac{2\nu}{2.3}(3,-3) + \frac{2\nu.3\nu}{2.3.4}(4,-4) - \ldots\right\} \\ &\quad - \frac{\nu}{2.3}\left\{\frac{1}{3}(3,-3) - \frac{3\nu}{3.4}(4,-4) + \frac{3\nu.4\nu}{3.4.5}(5,-5) - \ldots\right\} \\ &\quad + \frac{\nu^2}{3.4}\left\{\frac{1}{4}(4,-4) - \frac{4\nu}{4.5}(5,-5) + \frac{4\nu.5\nu}{4.5.6}(6,-6) - \ldots\right\} \\ &\quad - \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\end{aligned}$$

For $\nu=\frac{1}{2}$, $z=1$, $(2,-2)=4$, and $\nu^r(r,-r)=1$, so that, rearranging terms, since the series are absolutely and uniformly convergent:

$$\begin{aligned}&\underset{(\nu=\frac{1}{2},\ z=1)}{[\mathrm{X}_1]} \\ &= -4\mathrm{M}\left\{\frac{1}{2\cdot 1.2} - \frac{1}{3}\left(\frac{1}{1.2}+\frac{1}{2.3}\right) + \frac{1}{4}\left(\frac{1}{1.2}+\frac{1}{2.3}+\frac{1}{3.4}\right)\right. \\ &\qquad \left. - \frac{1}{5}\left(\frac{1}{1.2}+\frac{1}{2.3}+\frac{1}{3.4}+\frac{1}{4.5}\right) + \ldots\right\}\end{aligned}$$

$$= -4\mathrm{M}\left\{\frac{1}{2}\left(1-\frac{1}{2}\right)-\frac{1}{3}\left(1-\frac{1}{3}\right)+\frac{1}{4}\left(1-\frac{1}{4}\right)-\frac{1}{5}\left(1-\frac{1}{5}\right)+\ldots\right\}$$

$$= -4\mathrm{M}\left\{\frac{1}{2}-\frac{1}{3}+\frac{1}{4}-\frac{1}{5}+\ldots-\left(\frac{1}{2^2}-\frac{1}{3^2}+\frac{1}{4^2}-\frac{1}{5^2}+\ldots\right)\right\},$$

whence, using either

$$\sin x = x\prod_1^\infty\left(1-\frac{x^2}{r^2\pi^2}\right)$$

or a Tchebechev formula,

$$\underset{(\nu=\frac{1}{2},\ z=1)}{[\mathrm{X}_1]} = -4\mathrm{M}\left\{\frac{\pi^2}{12}-\log_e 2\right\} = -0{\cdot}51728\mathrm{M}.$$

XLVI. *The Effect of Lattice Distortion and Fine Grain on the X-Ray Spectra of Metals.* *By* W. A. Wood, *M.Sc., Physics Department, National Physical Laboratory, Teddington, Middlesex* *.

Summary.

An attempt is made to differentiate between the effects of fine grain and of lattice distortion on the X-ray spectrum of a metal. A method is suggested for calculating the proportionate contribution of each factor to the broadening of the X-ray diffraction lines. The results are applied to the particular case of electrodeposited nickel; it is shown that both factors must be taken into account in considering the observed diffusion of the X-ray spectra of nickel in this form.

Introduction.

A diffusion of the lines on the X-ray diffraction spectrum of a metal may be produced by the presence either of a very fine-grained structure or of lattice distortion. The first effect appears when the grain-size is about 10^{-4} cm. and grows rapidly more marked as the structure becomes still finer [1]. The second effect accompanies any permanent derangement of the atomic formation in the grains induced by excessive stresses [2]. The broadening normally is easily measurable and provides a useful

* Communicated by Dr. G. W. C. Kaye, O.B.E.

indication of the intrinsic structure of a material. But there arises the difficulty of discriminating between the two effects in any given case.

A distinct diffusion of the diffraction lines appears in the spectra of metals which have been subjected to cold-working. Also, an equally marked effect can occur in the spectra of metals such as chromium, iron, and nickel when produced by electrodeposition. The line-broadening in the case of cold-working is usually attributed to the distortion of the crystal lattice, but, in the electrodeposited metals, only to the production of a very fine grain. The latter assumption has been based on no very definite evidence, because hitherto no practical criterion has existed for distinguishing between the effects of distortion and grain-size. It is true that certain theoretical grounds for this purpose have been discussed [3] ; they depend on the fact that the relative breadths of the lines in different positions in a spectrum should indicate whether the broadening is due to one effect or the other. These considerations have, however, been complicated by the recent experimental observation that either effect can cause a selective anomalous broadening amongst the spectral lines [4] [5]. An alternative approach is suggested in the present paper ; it arises from work previously described which showed that as a metal is worked the distortion of the lattice rapidly grows to a maximum, characteristic of that metal [6]. This maximum will be referred to as the " lattice distortion limit " [7].

There is a limit, therefore, to the diffusion of a diffraction line as a result of cold-working a metal. But there is no such limit to the diffusion due to small grain-size. This idea is applied to the case of electrodeposited and cold-worked nickel with the aim of assigning the line-broadening, often observed in the spectrum of nickel plate, to its proper source.

Procedure.

The first series of observations was made to measure the change in breadth of the (311) line in the spectrum of a sample of nickel, the lattice of which was distorted to the limit. The second series was made to determine the breadth of the (311) line for a number of specimens

of nickel electrodeposited on copper. A final series was made to examine the effect on the same line of cold-rolling the electrodeposited specimens. The point of the last experiments is as follows. If the observed line-broadening is a consequence of the fine-structure only of the deposits, then the process of cold-working, by introducing lattice distortion as it does with ordinary nickel, should cause a further increase in breadth. An increase in breadth equivalent in amount to the lattice distortion limit would indicate that the original deposit was quite free from distortion ; on the other hand, no change in line-breadth would mean that the deposit was already distorted to the limit ; and an intermediate increase would connote the initial presence of partial distortion. In this way the proportionate contributions of lattice distortion and fine grain to the line-breadth might be inferred. We assume that the cold-work would not break down the grains to the extent of causing further line-broadening on that particular account. This assumption is justified by the existence of the observed upper limit to the broadening which can accompany cold-work.

Experimental Data.

(i.) *Electrodeposition.*—The nickel was deposited on strip copper from a bath with the following composition : 300 gm. of nickel sulphate, 6 gm. of boracic acid, and 3 gm. of sodium chloride per litre of distilled water. The anode was a rod of electrolytic nickel. The copper cathode received a prior treatment for cleaning the surface ; this involved first a wash in benzene and alcohol, next, cathodic bombardment in a solution of 50 gm. of sodium carbonate, 5 gm. of caustic soda, and 5 gm. of sodium cyanide per litre, and, finally, a short immersion in nitric acid followed by a rinse in distilled water. The deposition was carried out at 20° C. with a current density which varied, for different specimens, from 3 to 70 milliamps. per sq. cm. The time of deposition was adjusted so as to get deposits in each case of the same order of thickness, 0·1 mm. This value greatly exceeds that at which line-diffusion might occur on account of a thin coating. It was found that the width of the X-ray spectral lines from the different specimens varied between comparatively wide limits according to the current

density with which they were secured. In general a very high and a very low current density gave a deposit exhibiting a sharp spectrum; and an intermediate value to a spectrum definitely diffused. This last type of spectrum was usually associated with the harder kinds of coating, a point receiving further attention.

(ii.) *Cold-working.*—A strip of electrolytic nickel was annealed *in vacuo* and then cold-rolled by twelve successive stages from a thickness of 1·02 mm. to 0·27 mm. An X-ray diffraction photograph was secured at each stage under standard geometrical conditions, so that changes in the state of the spectrum could only be caused by changes in the structure of the specimen. The width of the (331) lines was measured at each step with the aid of a Moll microphotometer, and was plotted against the reduction in thickness of the strip. In this way the maximum width of line obtainable as a result of cold-working was secured. Fuller details of this procedure have been published previously [7]. A similar treatment involving reductions up to about 40 per cent. of the original thickness was given to the specimens of nickel-plated copper strip. The width of the (311) line, photographed under the standard conditions, was compared before and after rolling the specimen.

(iii.) *X-Ray Data.*—Each specimen to be photographed was mounted similarly at the centre of a cylindrical camera of radius 5·5 cm. to which the X-ray beam was directed by means of a collimator slit of length 4 cm. and diameter 1 mm. The characteristic Kα radiation of iron was used; with this wave-length the (311) line occurred at a glancing angle of 66·3° and, under normal conditions, gave the $\alpha\alpha_1$ doublet sharply with a separation of the components amounting to about 0·4 mm. As lattice distortion or small grain-size occurred, the doublets diffused into a single broad line. (It would have been possible by using copper radiation to have examined lines occurring at larger glancing angles and showing greater dispersion. But it was found difficult to avoid an enhanced background in photographs of the cold-worked specimens.) The exposures were reduced to a minimum consistent with ease of measurement in order to obviate the spreading of a line so often produced in a film by prolonged exposure

to X-rays. An attempt was also made to regulate the exposures so that the intensity of the lines in each spectrum were approximately the same. The after treatment of the films was conducted under constant conditions of time and temperature.

Observations.

The breadths of the (311) line for the specimens of rolled nickel are given in Table I. It is possible to measure the breadth along the line forming the background of radiation at the base of the peak as it appears on the microphotometer records, or at a point on the peak where the intensity is a defined fraction of that at the maximum. It was considered sufficient for purposes of comparison

TABLE I.

Times rolled.	Thickness.	Line-breadth of (311).
	mm.	mm.
1	0·98	2
2	0·90	4
3	0·80	6
4	0·72	8
5	0·65	8
6	0·59	9
7	0·53	7
8	0·48	7
9	0·43	8
10	0·38	8
11	0·33	7
12	0·27	7

to take the breadth of the line where the intensity was a half the maximum. A correction was made for presence of the α_1 component of the $\alpha\alpha_1$ line. The formula calculated by Brill [8] was employed, from which we have

$$b = \frac{B}{2}\left\{1+\sqrt{1-\frac{4\delta}{3B}}\right\},$$

where B is the measured breadth, b the corrected value, and δ the separation of the doublet as calculated for the conditions of working. The third column of the table therefore gives the breadths as defined above and as measured on the photographic record of the Moll microphotometer ; this instrument was geared to give a magnification of seven times.

It is seen from the table that after the fourth reduction the breadth of the (311) line ceases to increase and becomes practically constant at a value which averages out at 8 mm. The condition of the line at this stage is reproduced from the microphotometer record in fig. 1.

The breadths of the (311) lines determined in the same way, for a series of electrodeposited samples, are given in Table II. It is seen that the deposits give a line varying in breadth from 3 to 11 mm., as measured by the microphotometer.

Fig. 1.

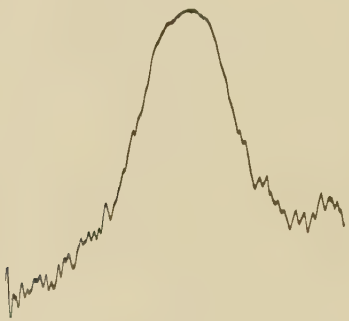
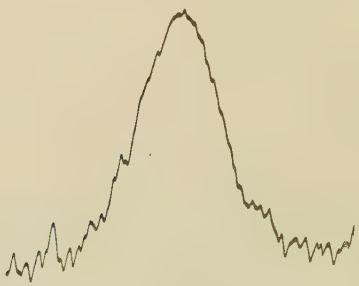


Fig. 2.



The effect of cold-working, also shown in the table, is best considered by noting separately (*a*) deposits from which the (311) line is less in width than the broadest line, 8 mm., obtainable by cold-working ordinary nickel, and (*b*) those for which the line is definitely wider than the lattice distortion maximum. It is then seen that in the case of the first type, the effect of cold-working was invariably to produce a large increase in the line-breadth. This is illustrated by figs. 2 and 3 ; the line-breadth before

rolling is 7 mm. and after 13 mm., showing an increase of 6 mm. On the other hand, in the case of type (*b*),

Fig. 3.

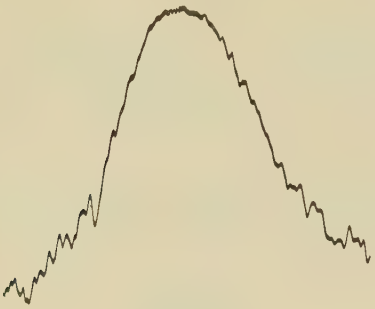


Fig. 4.

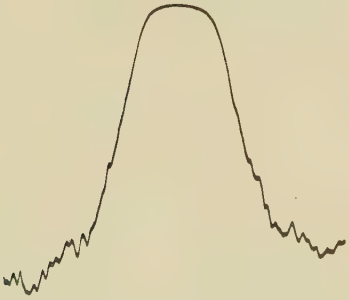
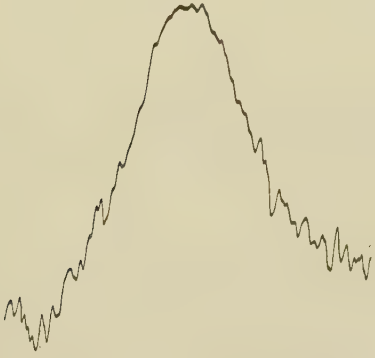


Fig. 5.



the result of rolling was to cause little or no change in the line-breadth. This is shown by figs. 4 and 5 ; the breadth of the line in each instance is 11 mm.

Conclusions.

We draw the following inferences from the observations :—

(1) The diffusion of an X-ray spectrum of electro-deposited nickel may in some circumstances be due to the occurrence of an ultra-fine grain. This follows from the fact that the line-breadth, as in the case of specimen (i.), can exceed the limiting value associated with lattice-distortion.

(2) The fine-grained structure of nickel deposits can be accompanied by lattice distortion. This is to be inferred

Table II.

Specimen.	(311) line before rolling.	Rolling reduction.	(311) line after rolling.
	mm.	per cent.	mm.
a.....	3	28	8
b	5	30	8
c	6	30	9
d.....	7*	43	13†
e	9	29	11
f	10	34	11
g	10	30	9
h	11	22	10
i	11**	40	11††

*Fig. 2. †Fig. 3.
**Fig. 4. ††Fig. 5.

from the observation that in a number of specimens the process of further working produces little or no change in the line-breadth, and that in no instance did the working cause an increase equal in degree to the distortion limit.

(3) The contribution of lattice distortion to the line-breadth may be a very appreciable proportion of the whole. The concept of a lattice-distortion limit can be applied to the calculation of the separate effects of distortion and fine-grain. If β is the breadth of a line as measured, B the contribution due to distortion, and b that due to grain-size, then before cold-working

$$\beta = \mathrm{B} + b.$$

After cold-working, when the lattice distortion gives its limiting effects, B_0, the line will broaden to the value β_1 where

$$\beta_1 = B_0 + b.$$

The value of B_0 being known from measurements on ordinary nickel, we have, in terms of known quantities,

$$b = \beta_1 - B_0$$

and

$$B = \beta - (\beta_1 - B_0).$$

Thus in the case of specimen (*d*), we find $b=5$ mm. and $B=2$ mm.; in specimen (i.), $b=3$ mm. and $B=8$ mm. These results show that no one effect can be regarded to the total exclusion of the other.

It is assumed, again, that the value of b is not increased by rolling in accord with the observation that B does not increase indefinitely on cold-working. The opposite tendency might well be the case, where the grain-size grows on working by re-crystallization; but this would not affect the main conclusions.

Remarks.

If the lines of an X-ray spectrum have a breadth greater than the value corresponding to the latitce-distortion limit, then it can be safely concluded that the broadening is due in some degree to a very fine-grained structure. If, however, the lines are less broad than the distortion limit, then further experiments are necessary to ascertain the relevant cause. The type of procedure suggested is that of subjecting the material to cold-work and examining the effect on the breadth of a given line. If the line does not alter lattice distortion is present, and any excess of breadth over the known distortion limit gives the contribution due to grain-size. If the line changes in breadth then the proportionate effects of distortion and grain-size may be estimated in the manner described. The only case where the effect of lattice distortion could be neglected would be when the line-breadths are very large in comparison with the value at the lattice-distortion limit Chromium plate affords an instance of this type. In other cases of fine grain in electrodeposited and other forms of metals, calculations of grain-size by the Laue equations should take into account the possibility of

corrections due to the presence of lattice distortion. A neglect of this precaution may have led in many cases to results of doubtful validity.

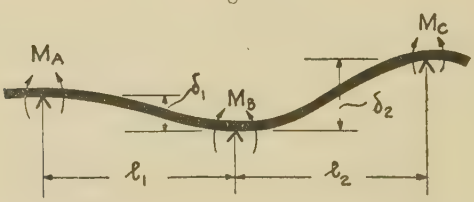
References.

(1) M. v. Laue, *Z. Krist*, liv. p. 115 (1926).
(2) V. Dehlinger, *Z. Krist.* lv. p. 615 (1927).
(3) V. Dehlinger, *Z. Metallkunde*, xxiii. p. 147 (1931).
(4) W. A. Wood, ' Nature,' cxxvii. p. 703 (1931).
(5) W. A. Wood, ' Nature,' cxxix. p. 760 (1932).
(6) W. A. Wood, Proc. Phys. Soc. xliv. p. 67 (1932).
(7) W. A. Wood, Phil. Mag. xiv. p. 656 (1932).
(8) R. Brill, *Z. Krist.* lxviii. p. 387 (1928).

XLVII. *Some Applications of the Theorem of Three Moments.* *By* V. Belfield, *M.A.**

THE object of the following paper is to bring to notice some applications of the Theorem of Three Moments in solving problems relating to the stresses in

Fig. 1.



certain types of framework, particularly those which contain rigid joints. Such problems can be solved in a straightforward though laborious manner by means of the well-known theory of bending for beams; they can also be solved † by consideration of the strain energy due to bending. The following method is perhaps of interest in that the equations to be solved are of greater generality, and moreover can be written down at once.

The form in which the Theorem of Three Moments for a beam will be used is as follows (the notation will be clear from fig. 1) :—

* Communicated by Prof. R. V. Southwell, M.A., F.R.S.
† C. E. Larard, M.Inst.C.E., Phil. Mag. xi. (May 1931).

$$\frac{1}{6}\left[M_A\frac{l_1}{I_1}+M_C\frac{l_2}{I_2}+2M_B\left(\frac{l_1}{I_1}+\frac{l_2}{I_2}\right)\right]$$
$$=\frac{A_1\bar{x}_1}{l_1 I_1}+\frac{A_2\bar{x}_2}{l_2 I_2}-\frac{E\delta_1}{l_1}-\frac{E\delta_2}{l_2}, \quad . \quad (1)$$

where

M_A, M_B, M_C are the hogging B.M.'s at the supports A, B, C respectively, the supports being either real or *imagined for convenience.*

l_1, l_2 are the distances between the supports.

A_1, A_2 are the areas of the B.M. diagrams, for the two spans l_1, l_2 respectively, due to the load carried, each length being considered as a separate span resting freely on supports at its ends.

$\bar{x}_1$ is the distance of the C.G. of the area A_1, measured from A.

$\bar{x}_2$ is the distance of the C.G. of the area A_2, measured from C.

δ_1 is the deflexion of the support B relative to A, reckoned positive when A is above B.

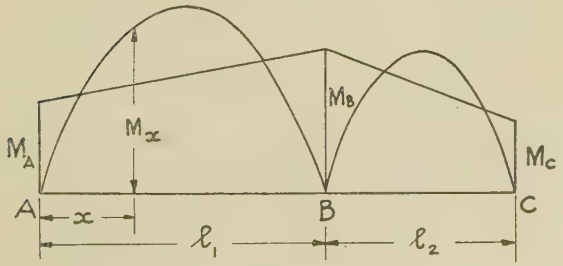
δ_2 is the deflexion of the support B relative to C, reckoned positive when C is above B.

Proof of the Theorem of Three Moments.

2. A short proof of the above theorem now follows :

Fig. 1 *a*.

Fig. 1 *b*.



Consider the portion AB. of the beam shown in fig. 1 *a*.

Taking A as origin of coordinates for the centre line of the beam, x being measured to the right and y downwards,

$$-\mathrm{EI}_1\frac{d^2y}{dx^2} = \text{sagging bending moment},$$
$$= \mathrm{M}_x - \left\{\mathrm{M}_\mathrm{A} + \frac{x}{l_1}(\mathrm{M}_\mathrm{B} - \mathrm{M}_\mathrm{A})\right\}, \quad . \quad (1\,a)$$

in the notation of fig. 1 *b*.

Multiplying this equation by x and integrating from $x=0$ to $x=l_1$, we have

$$-\mathrm{EI}_1\left[x\frac{dy}{dx} - y\right]_{x=0}^{x=l_1} = \int_0^{l_1}\mathrm{M}_x x\,dx$$
$$-\left\{\mathrm{M}_\mathrm{A}\frac{l_1^2}{2} + (\mathrm{M}_\mathrm{B} - \mathrm{M}_\mathrm{A})\frac{l_1^2}{3}\right\}. \quad . \quad (1\,b)$$

If i is the slope of the beam at B, this becomes, in the notation of equation (1)

$$-\mathrm{EI}_1[l_1 i - \delta_1] = \mathrm{A}_1\bar{x}_1 - \left\{\mathrm{M}_\mathrm{A}\frac{l_1^2}{2} + (\mathrm{M}_\mathrm{B} - \mathrm{M}_\mathrm{A})\frac{l_1^2}{3}\right\}. \quad (1\,c)$$

Similarly for the portion CB, with origin at C,

$$-\mathrm{EI}_2[-l_2 i - \delta_2] = \mathrm{A}_2\bar{x}_2 - \left\{\mathrm{M}_\mathrm{C}\frac{l_2^2}{2} + (\mathrm{M}_\mathrm{B} - \mathrm{M}_\mathrm{C})\frac{l_2^2}{3}\right\}. \quad (1\,d)$$

Dividing (1*c*) by $l_1\mathrm{I}_1$, and (1*d*) by $l_2\mathrm{I}_2$, and adding, we have

$$\frac{1}{6}\left[\mathrm{M}_\mathrm{A}\frac{l_1}{\mathrm{I}_1} + \mathrm{M}_\mathrm{C}\frac{l_2}{\mathrm{I}_2} + 2\mathrm{M}_\mathrm{B}\left(\frac{l_1}{\mathrm{I}_1} + \frac{l_2}{\mathrm{I}_2}\right)\right]$$
$$= \frac{\mathrm{A}_1\bar{x}_1}{l_1\mathrm{I}_1} + \frac{\mathrm{A}_2\bar{x}_2}{l_2\mathrm{I}_2} - \frac{\mathrm{E}\delta_1}{l_1} - \frac{\mathrm{E}\delta_2}{l_2} \quad . \quad (1)$$

The formula requires modification when E is different for the two portions of the beam, and then becomes

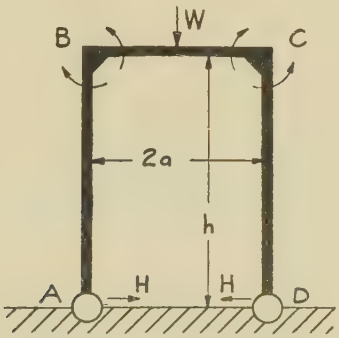
$$\frac{1}{6}\left[\mathrm{M}_\mathrm{A}\frac{l_1}{\mathrm{E}_1\mathrm{I}_1} + \mathrm{M}_\mathrm{C}\frac{l_2}{\mathrm{E}_2\mathrm{I}_2} + 2\mathrm{M}_\mathrm{B}\left(\frac{l_1}{\mathrm{E}_1\mathrm{I}_1} + \frac{l_2}{\mathrm{E}_2\mathrm{I}_2}\right)\right]$$
$$= \frac{\mathrm{A}_1\bar{x}_1}{l_1\mathrm{E}_1\mathrm{I}_1} + \frac{\mathrm{A}_2\bar{x}_2}{l_2\mathrm{E}_2\mathrm{I}_2} - \frac{\delta_1}{l_1} - \frac{\delta_2}{l_2}. \quad (2)$$

On the other hand, when I is the same for the two spans considered, the formula simplifies to

$$\frac{1}{6}\left[\mathrm{M}_\mathrm{A}l_1 + \mathrm{M}_\mathrm{C}l_2 + 2\mathrm{M}_\mathrm{B}(l_1 + l_2)\right]$$
$$= \frac{\mathrm{A}_1\bar{x}_1}{l_1} + \frac{\mathrm{A}_2\bar{x}_2}{l_2} - \frac{\mathrm{EI}\delta_1}{l_1} - \frac{\mathrm{EI}\delta_2}{l_2} \quad . \quad (3)$$

Whichever form of the equation is entailed, the unknowns are usually one, or all, of M_A, M_B, M_C, δ_1 and δ_2, δ_1 and δ_2 being zero if the supports are at the same level. If a uniformly distributed load w_1 per foot run is carried by the span l_1, then $\frac{A_1\bar{x}_1}{l_1}$ is seen to be $\frac{w_1 l_1^3}{24}$; a single concentrated load W_1 at the centre of the span l_1 gives the value $\frac{W_1 l_1^2}{16}$, and if W_1 is not at the centre of the span the value of $\frac{A_1\bar{x}_1}{l_1}$ is easily determined, since the B.M. diagram consists of two triangles, so that the product $A_1\bar{x}_1$ can be immediately written down as the sum of the separate products for each triangle.

Fig. 2.



Application to the "Portal Frame Structure."

3. Fig. 2 represents a simple type of "portal" structure supporting a weight W centrally placed, the joints at the ground being hinged and those at the top being assumed rigid. In this and similar problems the shortening effect of the direct thrust in the members will be neglected; in considering the strain energy the corresponding simplification would be to neglect that due to direct thrust as compared with that due to bending; then, from the symmetry of the figure and of the loading, it follows that the points B and C will not move either vertically or horizontally. We shall assume that E and I have the same values throughout the structure.

We now imagine the structure "unwrapped" and opened out into a "straight" continuous beam with imaginary supports at A, B, C, and D, which, in this case, will all be

at the same level (see fig. 3). It will be clear that the sole purpose of performing this apparently unjustifiable operation on the structure and of using fig. 3 as our "working" diagram for applying the Three Moments equation is merely to avoid turning our paper—and our mind with it—through two separate right angles, as we consider the relative displacement of the ends of the different portions of the beam. We have, of course, still preserved the rigidity of the joints B and C, since we are considering the beam to be "continuous" at these points.

Fig. 2 *a*.

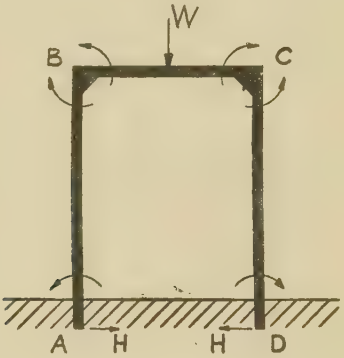
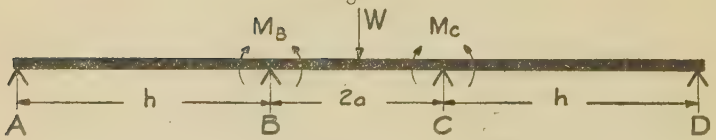


Fig. 3.



Applying equation (3) to the portion ABC, and noticing that

$$M_A=0\,;\quad M_B=M_C\,;\quad A_1x_1=0\,;\quad \delta_1=\delta_2=0\,;$$

we derive the equation

$$\frac{1}{6}\left[M_B(6a+2h)\right]=\frac{Wa^2}{4}\,;\quad .\quad .\quad .\quad .\quad (4)$$

whence

$$M_B=\frac{3}{4}W\frac{a^2}{(3a+h)}\,;\quad .\quad .\quad .\quad .\quad (5)$$

The bending moment diagram for each portion of the structure can now be determined, and hence the stress at

any point, because statical considerations show that the horizontal reaction at A or D (fig. 2) is given by

$$H = \frac{M_B}{AB} = \frac{3}{4} W \frac{a^2}{h(3a+h)}. \quad . \quad . \quad . \quad (6)$$

If the ends A and D of the structure are encastred, as in fig. 2 *a*, the process of "unwrapping" leads to fig. 4, the conditions at the two ends A and D being reproduced by two "supports" at each of these points, indefinitely close

Fig. 4.

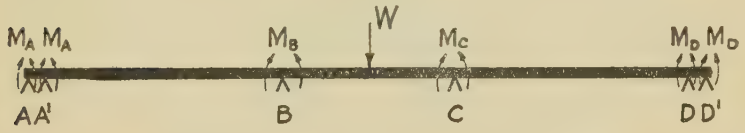
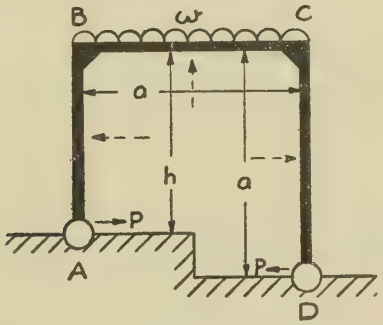


Fig. 5.



together. We now apply equation (3) to the portions AA'B, A'BC, in turn. Putting $M_C = M_B$, we obtain

$$M_B h + 2M_A h = 0, \quad . \quad . \quad . \quad . \quad . \quad (7)$$

and

$$\frac{1}{6}\left[M_A h + M_B 2a + 2M_B(2a+h)\right] = \frac{W a^2}{4}, \quad . \quad (8)$$

whence

$$M_A = M_D = -\frac{M_B}{2} = -\frac{1}{2} W \frac{a^2}{(4a+h)}. \quad . \quad . \quad (9)$$

The horizontal reaction at A (fig. 2 *a*) is H, where

$$H \times h = M_B - M_A. \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Hence

$$H = \frac{3}{2} W \frac{a^2}{h(4a+h)}. \quad . \quad . \quad . \quad . \quad (11)$$

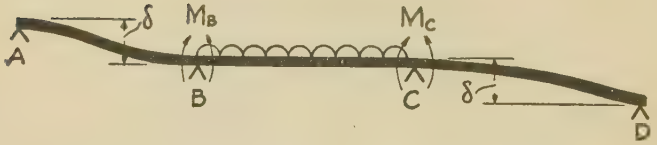
The above examples serve to illustrate cases in which the position of the load W is such as to introduce complete symmetry. If the load is not at the centre of the portion BC, the structure will deflect to the right or left, the points B and C each moving horizontally, say, by an amount δ.

Consider the case shown in fig. 5, where the two vertical portions differ in length, and the top member carries a uniformly distributed load of w per ft. run. Assuming that the structure will deflect to the right, B and C each moving by an amount δ, the "unwrapped" diagram (fig. 6) is obtained by viewing the structure as shown by the dotted arrows in fig. 5. Noticing that $M_A = M_D = 0$, we now write down the following equations for the portions ABC, BCD respectively:

$$\frac{1}{6}\left[M_C a + 2M_B(a+h)\right] = \frac{wa^3}{24} - \frac{EI\delta}{h}, \quad . \quad . \quad (12)$$

$$\frac{1}{6}\left[M_B a + 2M_C \cdot 2a\right] = \frac{wa^3}{24} + \frac{EI\delta}{a}. \quad . \quad . \quad (13)$$

Fig. 6.



If P is the horizontal reaction at A and at D (fig. 5), we have

$$M_B = Ph, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

$$M_C = Pa, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

and hence obtain, on eliminating δ,

$$P = \frac{wa^3}{8} \cdot \frac{(a+h)}{(2a^3 + a^2h + ah^2 + h^3)}. \quad . \quad . \quad . \quad (16)$$

Equations (14), (15) and (16) provide all that is required for determining the bending moment diagram, and the value of δ can be found by substitution in (12) or (13).

4. A more general type * of "gallows" structure with unsymmetrical loading is shown in fig. 7, the left-hand vertical member being subjected to a uniformly distributed horizontal load of intensity p and the top member to a loading of intensity q. The moments of inertia of the left-hand,

* This is C. E. Larard's problem (see footnote on p. 563).

top and right-hand portions will be taken to be nI, mI and I respectively.

The "unwrapped" diagram is shown in fig. 8. Taking Young's Modulus to be the same for each portion, we write down four equations of type (1) ; denoting the horizontal reaction at A by P, and that at D by $(hp-P)$, we obtain two more equations by considering the equilibrium of the two vertical portions separately. These six equations determine M_A, M_B, M_C, M_D, P, and δ.

Fig. 7.

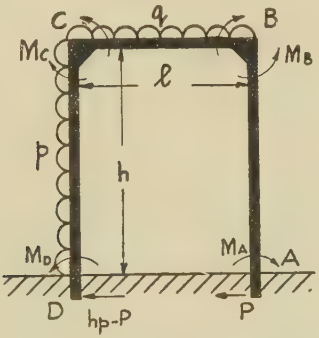


Fig. 8.

Writing

$$\alpha=\frac{ph^2}{4}; \quad \beta=\frac{ql^2}{4}; \quad \gamma=\frac{6EI\delta}{h^2},$$

we obtain for the first four equations :

$$M_C+2M_D=\alpha+n\gamma, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

$$M_Dmh+M_Bnl+2M_C(mh+nl)=mh\alpha-nl\beta-mnh\gamma. \quad (18)$$

$$M_Cl+M_Amh-2M_B(mh+l)=l\beta+mh\gamma, \quad . \quad . \quad . \quad . \quad (19)$$

and

$$M_B+2M_A=-\gamma, \quad . \quad . \quad . \quad . \quad . \quad (20)$$

and for the other two :

$$M_A-M_B=-Ph, \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

$$M_D-M_C=-Ph+2\alpha. \quad . \quad . \quad . \quad . \quad (22)$$

These six equations reduce quickly to three simultaneous equations for, say, M_A, M_B, M_C, the full results being as follows:

$$M_A = -\frac{1}{4D}\Big[\{18m^2h^2+(22n+13)mhl+9l^2n\}ph^2 + \{3(5n-1)mh+n(n+1)l\}ql^3\Big]; \quad . \quad (23)$$

$$M_B = +\frac{1}{4D}\Big[\{18m^2h^2+(26n-1)mhl\}ph^2 + \{3(7n+1)mh+2n(n+1)l\}ql^3\Big]; \quad . \quad (24)$$

$$M_C = -\frac{1}{4D}\Big[\{6(2n-1)m^2h^2+(25n-2)mhl\}ph^2 - \{3n(n+7)mh+2n(n+1)l\}ql^3\Big]; \quad . \quad (25)$$

$$M_D = +\frac{1}{4D}\Big[\{6(4n+1)m^2h^2+(12n^2+59n+2)mhl + 3(4n+1)nl^2\}ph^2+\{3n(n-5)mh - n(n+1)l\}ql^3\Big]; \quad . \quad (26)$$

$$Ph = \frac{3}{4D}\Big[\{12m^2h^2+4(4n+1)mhl+3nl^2\}ph^2 + \{12mnh+n(n+1)l\}ql^3\Big]; \quad . \quad . \quad (27)$$

$$\delta = \frac{3h^2}{8EI}\cdot\frac{1}{D}\Big[\{2m^2h^2+(2n+3)mhl+2nl^2\}ph^2 + \{(n-1)mh\}ql^3\Big]; \quad . \quad (28)$$

where D is

$$6\Big[(n+1)\{3m^2h^2+nl^2+(n+1)mhl\}+9mnhl\Big]. \quad (29)$$

Special Cases.

The above results may be readily used for deducing the solution in special cases, such as:

(*a*) n or m, or both, being equal to unity.

(*b*) p or q being zero.

If pin joints are assumed at A and D the problem is best attacked afresh, the resulting equations being very much less complicated.

5. The methods used above can be extended to the calculation of stresses in a framework such as is shown in fig. 9, where the joints at A and B are assumed rigid and the ends are encastred.

Let the bending moments in the different portions at the points A, C, and D be as shown. Considering equilibrium of the joint A, we have

$$2M_2 + M_3 = 0. \quad . \quad . \quad , \quad . \quad . \quad (30)$$

For the portion DAB we obtain fig. 10 as the "unwrapped" diagram. It will be noticed that at the point A we have

Fig. 9.

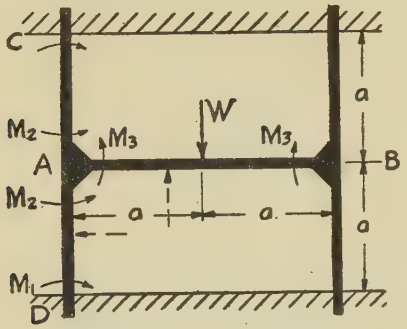
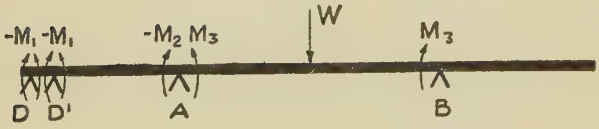


Fig. 10.



two bending moments to consider: particular attention to sign is necessary, since in equation (1), (2) or (3) the bending moments must be written down as positive when "hogging." It will be seen (fig. 10) that we have a discontinuity in the value of the bending moment at A by an amount M_2, which is the bending moment at that point in the portion AC. If reference is now made to the bending moment diagram (fig. 1 *b*) used in the proof of the Theorem of Three Moments, it will be clear that equations (1 *c*) and (1 *d*) will still hold for the portion D'AB in fig. 10, where M_B in equation (1 *c*) is in this case $-M_2$, and M_B in equation

$(1d)$ is $+M_3$; hence the last term of the left-hand side of equation (1) becomes

$$2(-M_2)a+2.M_3.2a.$$

Hence for the portion D′AB we derive the equation

$$\frac{1}{6}\Big[-M_1a+M_3.2a+2(-M_2)a+2M_3.2a\Big]=\frac{Wa^2}{4}. \quad (31)$$

Again, for the portion DD′A (fig. 10), or for CC′A, we have

$$\frac{1}{6}\Big[-M_2a+(-2M_1)a\Big]=0. \quad . \quad . \quad . \quad (32)$$

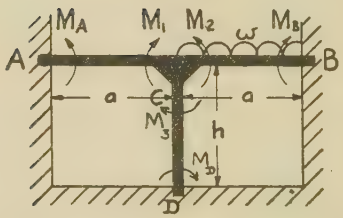
From (30), (31) and (32), we have at once

$$M_1=\frac{Wa}{18}; \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (33)$$

$$M_2=-\frac{Wa}{9}; \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (34)$$

$$M_3=\frac{2Wa}{9}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (35)$$

Fig. 11.



A similar problem is illustrated in fig. 11, the point C being assumed to remain stationary ; the results are given below, where I, I_1 and I_2 are the moments of inertia of the portions AC, BC, and CD respectively, and $\lambda=\frac{h}{a}$, $\mu=\frac{I_1}{I_2}$:

$$M_A=-\frac{wa^2}{24}\cdot\frac{\lambda\mu}{1+2\lambda\mu};$$

$$M_B=\frac{wa^2}{24}\cdot\frac{2+5\lambda\mu}{1+2\lambda\mu};$$

$$M_D=\frac{wa^2}{24}\cdot\frac{1}{1+2\lambda\mu};$$

$$M_1 = \frac{wa^2}{12} \cdot \frac{\lambda\mu}{1+2\lambda\mu};$$

$$M_2 = \frac{wa^2}{12} \cdot \frac{1+\lambda\mu}{1+2\lambda\mu};$$

$$M_3 = \frac{wa^2}{12} \cdot \frac{1}{1+2\lambda\mu}.$$

6. *Application to find the Deflexion of Beams under simple types of loading.*

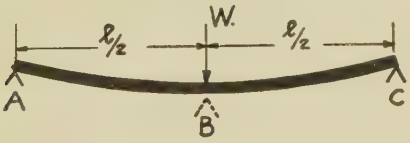
A horizontal beam simply supported at its ends and carrying a concentrated load W at its centre is shown in fig. 12. We notice that in equation (3)

$$M_A = M_C = 0; \quad M_B = -\frac{Wl}{4};$$

$$l_1 = l_2 = \frac{l}{2};$$

and $\delta_1 = \delta_2 = \delta$, say, the central deflexion.

Fig. 12.



Hence we write down

$$\frac{1}{6} \times 2\left(-\frac{Wl}{4}\right) l = -2\frac{EI\delta}{l/2},$$

whence

$$\delta = \frac{1}{48}\frac{Wl^3}{EI}.$$

If the beam carries a uniformly distributed load of intensity w per ft. run, we write down the equation for the central deflexion as follows :

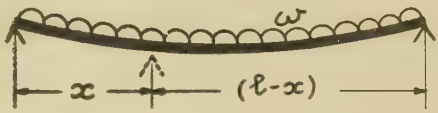
$$\frac{1}{6} \times 2\left(\frac{-wl^2}{8}\right) l = 2\frac{w\left(\frac{l}{2}\right)^3}{24} - 2\frac{EI\delta}{l/2},$$

whence

$$\delta = \frac{5}{384}\frac{wl}{EI}.$$

In the above case (fig. 13) if we require to know the

Fig. 13.



deflexion at a point distant x from the left-hand end, we derive the equation below, since

$$M_x = -\frac{wx}{2}(l-x) :—$$

$$-\frac{1}{6}2\frac{wx}{2}(l-x)l = \frac{w}{24}\{x^3+(l-x)^3\} - EI\delta\left(\frac{1}{x}+\frac{1}{l-x}\right),$$

whence $$\delta = \frac{w}{24EI}x(l-x)(l^2+lx-x^2).$$

XLVIII. *On the Continuous Spectrum of Sodium.* *By* H. Hamada*. (From the Laboratory of Physics, Sendai, Japan.)

[Plate XIII.]

Introduction.

A number of observers have already reported that the vapour of alkali metals emits a continuous spectrum in the flame. Gouy †, afterwards Lenard ‡, considers that this continuous spectrum consists of two distinct spectra; one being formed of the haloes of the principal series lines and having intense and broad maxima about these lines, and the other consisting of the haloes of the subordinate series lines, covering a very wide range of wave-lengths uniformly and forming a continuous background. Leder § has measured the intensity distribution of this continuous spectrum, for the purpose of confirming Lenard's view. The continuous background has been observed by Wiedemann and Schmidt ‖, Kalähne ¶, and also by Gehlhoff and

* Communicated by Prof. J. Okubo.
† M. Gouy, *Ann. de Chim. et Phys.* (5) xviii. p. 1 (1879).
‡ P. Lenard, *Ann. d. Phys.* xvii. p. 197 (1905).
§ F. Leder, *Ann. d. Phys.* xxiv. p. 305 (1907).
‖ E. Wiedemann and G. Schmidt, Wied. *Ann.* lvii. p. 451 (1896).
¶ A. Kalähne, Wied. *Ann.* lxv. p. 829 (1898).

Rottgardt * in the positive column of the glow discharge through sodium vapour in a condition of comparatively high pressure. On account of this experimental result, and also because this background is emitted more intensely in the flame of oxyhydrogen gas than in that of coal gas, and emitted intensely in the region of the flame, where the atomic ions exist in large concentration, and for the further reason that its intensity varies nearly proportionally to the degree of ionization in the vapour, viz., its intensity increases in the order of lithium, sodium, potassium, rubidium, and cæsium, Stark † has explained it as being due to the recombinations of atomic ions with electrons.

In spite of the fact that the continuous spectrum, which is described as the halo of the principal series line, has long been observed, it is to be regretted that as yet its intensity distribution and wave-length limits have not been precisely measured, and that its origin is not yet known. Kuhn ‡ has recently observed that, in the absorption spectra of sodium, potassium, and cæsium vapours at high pressure, the narrow bands appear in the vicinity of the lines of the principal series. He considers that the band is emitted from the atomic molecules loosely combined by the polarization force. Pringsheim and Jablonski §‖ have also found that in the fluorescence of sodium vapour at high pressure excited by a white light, the very broadened D-lines are emitted together with the ordinary resonance radiation, and they have also explained that the absorption and emission of these broadened D-lines are due to the polarization molecules, above mentioned. Of course, the narrow band (or the broadened D-lines) is observable only in the neighbourhood of the atomic lines, so it appears that this band is different from the continuous spectrum extending in the wide range, such as that observable in the flame. But, from the results of this experiment, it seems that they are not quite independent, and that there is some intimate connexion between the two.

It has already been pointed out in the previous ¶ report that the diatomic molecules of mercury, cadmium, zinc, and magnesium emit the discrete or diffused band series accom-

* G. Gehlhoff and K. Rottgardt, *Verh. d. D. Physik. Ges.* xii . p. 492 (1910).

† J. Stark, *Ann. d. Phys.* lii. p. 255 (1917).

‡ H. Kuhn, *Naturwiss.* xviii. p 332 (1930); *Zs. f. Phys.* lxxvi. p. 782 (1932).

§ P. Pringsheim, 'Fluorescenz und Phosphorescenz,' p. 73.

‖ A. Jablonski and P. Pringsheim, *Zs. f. Phys.* lxx. p, 593 (1931); lxxiii. p. 281 (1932).

¶ H. Hamada, Phil. Mag. xii. p. 50 (1931).

panied by the continuous spectrum about the resonance line, and that this band series appears on the longer wave-length side of that line, the convergence limit being near the line, and further that the series suddenly breaks off or melts into the continuous spectrum on that side, while the latter extends in a wide region from the maximum at the resonance line, especially on the longer wave-lengths side. It was also stated that the discrete or diffused band series is emitted by the stable excited molecules during the transitions to the stable or unstable lower state, and the continuous spectrum about the resonance line by the excited atoms during their collisions with other normal atoms (quasi-molecules). By the experiment which is here dealt with, it is shown that from the results of the investigation of the intensity distribution of the continuous spectrum of sodium, which is generally considered to be the halo of the D-lines, and also of the relative intensity of it to the discrete band systems, under various conditions of excitation, the same explanation serves also to account for the emission of the molecular spectrum of sodium, as in the case of mercury, cadmium, zinc, and magnesium.

Experimental Arrangements.

For the purpose of concentrating the excited molecules of sodium, a hollow cathode of iron in a slightly modified form from Schüler's original design * was used. A thin and narrow tube having a small aperture at the end was attached to the main tube. in which the metals to be vaporized were placed in a suitable quantity, and it was contrived to heat the main tube and the attached tube separately. By this means, it was possible to vary the pressure and the temperature of the vapour satisfactorily, while emitting the light at the observing aperature. The temperature of the vapour at this aperture was measured with an optical pyrometer, or with a thermocouple of platinum and platinum-rhodium, and the amount of vapour pressure was indirectly measured from the temperature in the main tube. The spectrograph mainly used was a three-prism glass spectrograph with a strong illumination made by C. A. Steinheil Söhne & Co., and sometimes the second-order spectrum of a concave grating 21·5 feet in focal length was also observed. The photographic plates used were an Agfa panchromatic plate and an Eastmann infra-red sensitive plate. Owing to the marked reduction of the colour sensibility to light of a longer wave-length than about 7000 Å. of the former, and to the sensibility

* H. Schüler, *Zs. f. Phys.* xxxv. p. 323 (1926).

minimum in the red region of the latter plate, it was difficult to observe the intensity distributions of the band system and the continuous spectrum directly. Therefore, the spectra of sodium vapour emitted under various conditions and that of a tungsten lamp, the energy distribution of which was previously known, were photographed on the same plate with a nearly equal duration of exposure with Dorgelo's stepped reducer * with six steps, and the intensity distribution of the spectrum was calculated by the method used by him. The absorptive power of the stepped reducer was measured by the radiation which was obtained by passing the light from a tungsten lamp through a yellow filter (No. 12632 of Carl Zeiss), and a layer of aqueous solution of copper sulphate (57 gr. per litre) 1 cm. in thickness. Of course, the absorptive power of a platinum film will vary with the wave-length, but it will probably not differ conspicuously from the mean value, as above obtained.

Experimental Results.

It is well known that sodium molecules emit two band systems in the visible and infra-red region : one is generally designated the red-band and the other the green-band system, and both systems are observable in emission, in absorption, as well as in fluorescence.

(I.) Before describing the results of the observations in the spectral range where the red-band system mainly exists, it will be convenient to divide the spectrum into the following three kinds A, B, and C, according to the nature of the emission centres.

(A) The first kind is the ordinary red-band system, which is emitted by the transitions $^1\Sigma_u^+ \rightarrow {}^1\Sigma_g^+$ in the stable molecules, and appears in the whole region covering the wave-lengths from 8150 to 5500 Å. As the emission of this band system is mainly due to the excited molecules which are produced from the normal molecules in various vibrational states through collisions with electrons, or through the triple collisions, the bands emitted by the transitions from the various vibrational levels in the upper $^1\Sigma_u^+$ state are observable. As a consequence of the Franck-Condon principle, taking into consideration the distribution of the normal molecules in the various vibrational levels, it is to be expected that the intensity of this band system will be at its maximum in the longer wave-lengths region and will gradually fade away towards the violet.

* H. B. Dorgelo, *Phys. Zeit.* xxvi. p. 756 (1925).

(B) The second kind of spectrum is that which has already been observed by Wood and Galt * in their investigation of the "cathode ray fluorescence" of sodium vapour, and which has also been explained as the modification of the red-band system by Schüler †, Loomis and Nile ‡, and also by Kimura and Uchida §. This spectrum is emitted by the molecules excited to the special limited ranges of energy levels in the $^1\Sigma_u^+$ state, which result from collisions of the second kind between the normal molecules and the excited atoms in the lowest ^2P state, or in consequence of the absorption of the energy of the broadened D-lines by the normal molecules, and there are many intensity maxima ("resonance maxima") arranged on both sides of the D-lines. Therefore, the distances between the two neighbouring maxima are approximately those of lower vibrational levels in the normal molecule, and the intensities of these maxima decrease gradually with the increase of the wave-lengths until they disappear, and then again recur and increase towards the maximum at about 8000 Å., as shown in fig. 1 (Pl. XIII.), where the wave-number differences between the two maxima as measured are about 110 cm.$^{-1}$, which are smaller than those in the neighbourhood of the D-lines. The reason why the wave-number differences between the two maxima are small, and also why they appear in the narrower wave-length range with greater intensity than those near the D-lines, will be well understood by bearing in mind the fact that, as Loomis and Nile have stated, the probability of electronic transitions at the far-nuclear turning-points of vibration is greater than that at the near-nuclear turning-points, together with the fact that, in the final states of emission of the infra-red bands, the $^1\Sigma_g^+$ molecules have higher vibrational energies, and the potential curve of the molecule has a smaller gradient (dU/dr), in comparison with the final states of emission of the maxima about the D-lines. As the band system A and the resonance maxima B are both emitted from the transitions between the $^1\Sigma_u^+$ and the $^1\Sigma_g^+$ states of the stable molecules and have well-defined rotational structure, and also as their intensities frequently vary in agreement (though the greater the relative intensity of the latter, the higher the pressure and the lower the temperature of the vapour), they are frequently classified in one and the same

* R. W. Wood and R. H. Galt, Astrophys. J. xxxiii. p. 72 (1911).
† H. Schüler, *Zs. f. Phys.* xliii. p. 474 (1927).
‡ F. W. Loomis and S. W. Nile, Phys. Rev. xxxii. p. 873 (1928).
§ M. Kimura and Y. Uchida, Sci. Pap. Inst. Phys. Chem. Res. xviii. p. 119 (1932).

class in the following pages, and are designated the "discrete bands or band system."

(C) The third kind of spectrum is a continuous one which has an intensity maximum at the resonance lines and a sharp edge at about 5500 Å., while it fades away gradually towards the side of the longer wave-lengths (Pl. XIII. fig. 2). This spectrum is considered to be emitted mainly by the quasi-molecules, but may be partly emitted by the transitions from the stable states to the unstable lower states corresponding to the repulsive potential curve, above the horizontal asymptote, or by very loosely-combined molecules, and its character is markedly different from that of the discrete band system described in (A) and (B). Contrary to the case of the diatomic molecules of mercury, cadium, zinc, and magnesium —in which the discrete or diffused band system due to the stable molecules appears mainly in a wave-length region which is quite different from that in which the continuous spectrum due to quasi-molecules is emitted—it was observed that, in the case of the alkali metals, the ordinary band system of the discrete bands appeared superimposed on the continuous spectrum. Therefore, it is rather puzzling to observe the variation in the relative intensities of both kinds of spectrum, and that of the intensity distribution in each spectrum due to the change in exciting conditions. The observation of the intensity distribution of this continuous spectrum, therefore, can be accurately carried out only in the conditions in which the discrete bands do not appear.

(II.) The variations in intensity of the three kinds of spectrum due to the variation in the vapour pressure. First of all, the pressure being varied but without altering the current density and the temperature of the sodium vapour, the intensity both of the discrete band system and of the continuous spectrum was observed (Pl. XIII. fig. 3).

(1) In the case where the vapour was kept at a relatively low temperature ($\sim 500°$ C.), it was observed that when the vapour pressure is very small, namely, less than $\sim 0{\cdot}1$ mm., only the narrow atomic lines are emitted, but that when the pressure is gradually raised to ~ 1 or 2 mm. the breadths of D-lines become greater and greater, and the continuous spectrum C about the D-lines appears, in addition to the discrete bands A and B. As the pressure of the vapour is increased more and more, the intensity of the discrete bands as well as of the continuous spectrum are both increased considerably with this increase, and the increasing rate of the former is greater than that of the latter. As the intensity

of the continuous spectrum relative to the discrete bands is always small, the former is observable only in the neighbourhood of the D-lines on the photographic plate with the short duration of exposure.

(2) When the vapour is maintained at a high temperature ($\sim 1000°$ C.), and in the case where the vapour pressure is low, the atomic lines alone are intensely observed, and the continuous spectrum about the D-lines is absent or very weak, and the latter spectrum begins to enhance when the pressure was increased to $\sim$1 or 2 mm., and, with the further increase of pressure its intensity is markedly increased, along with the (apparent) broadening of the spectrum. At high vapour pressure ($\sim$6 or 7 mm.), the discrete bands A and B are developed and superimposed on the continuous spectrum, and the more intensely these bands, as well as the continuous spectrum appear, the higher the pressure of the vapour. It is also known that the increasing rate of the intensity of the former is greater than that of the latter, though it is not so conspicuous as in the case of a lower temperature.

Though the concentration of the excited molecules depends on various factors, it may be safely inferred that it varies with that of the normal molecules. As the concentration of the excited quasi-molecules is linearly proportional to the frequency of collisions between excited atoms and normal atoms, it may be approximately proportional to the vapour pressure, if we assume that the number of excited atoms is constant, independently of the pressure. Though the latter assumption is not realized in an actual case, the concentration of the excited quasi-molecules may not be so markedly increased, with the increase of vapour pressure, as in the case of the stable excited molecules, because the concentration of the normal molecules is proportional to the square of the vapour pressure. Therefore, it follows that, when the vapour pressure is increased, the discrete band system is more intensely enhanced than the continuous spectrum, the temperature being kept constant. The reason why this effect is more conspicuous at a lower than at a higher temperature may be explained by the fact that, at a low temperature, the probability of the occurrence of triple collisions, or of the collisions of the second kind, is more markedly increased with the increase of vapour pressure in comparison with the effect at a high temperature.

(III.) The intensity variations of the three kinds of spectrum due to the temperatures of the vapour. Looking

at the observed results, as given above, from another point of view, further knowledge regarding the intensity variations of the discrete band system and the continuous spectrum due to the temperature variation of the vapour is obtained.

(1) In the case where the vapour is kept at low pressure (~1 or 2 mm.) the discrete bands are mainly excited, while the continuous spectrum is relatively weak at the lower temperature (~500° C.). The rise of temperature causes, on the one hand, the relative enhancement of the continuous spectrum, and, on the other hand, a diffusing of the intensity maxima of the resonance series B and the fading away of the band-heads of the band system A. At a high temperature (~1000° C.), these discrete bands almost completely disappear, and the continuous spectrum alone is emitted.

(2) Similarly, it is known that in the case of higher pressure (~6 or 7 mm.), the more intensely the discrete bands A and B are emitted, compared with the continuous spectrum C, the lower the temperature of the vapour, but as the temperature of the vapour rises the case is reversed and the continuous spectrum becomes very intense and, therefore, broadens out widely towards both sides of the D-lines. It may be remarked here that these temperature effects are more conspicuously observed in the case of a higher than in that of a lower pressure, and also that the temperature, at which the continuous spectrum alone is excited with the fading away or disappearance of the discrete bands, depends on the vapour pressure; the higher this temperature the higher the vapour pressure.

In the case where the temperature of the vapour increases under a constant pressure, it is obvious that the concentration of normal atoms decreases inversely proportionally to the absolute temperature. Accordingly, it is to be expected that the intensity of the continuous spectrum due to the quasi-molecules will little decrease with the rise of temperature. But this rise causes a more rapid reduction of the concentration of the normal molecules than of the normal atoms, and, in addition, the probability of the triple collisions as well as of the collisions of the second kind, by which the excited molecules are produced, is considerably decreased, while the probability of the dissociation of the excited molecules by the collisions is increased by the rise of temperature. Therefore, at a high temperature, the intensities of the discrete bands A and B will be more considerably reduced than the continuous spectrum. Consequently, when the temperature is increased, the continuous spectrum is relatively enhanced against the discrete bands. The reason why

this effect is conspicuous at a high pressure may be understood from the reason stated above in the case of pressure variation. The concentration of excited molecules at high pressure is greater than that at low pressure, if the vapour is maintained at the same temperature, and in order to obtain the condition in which the continuous spectrum alone is excited the vapour must be heated to a higher temperature in the former case.

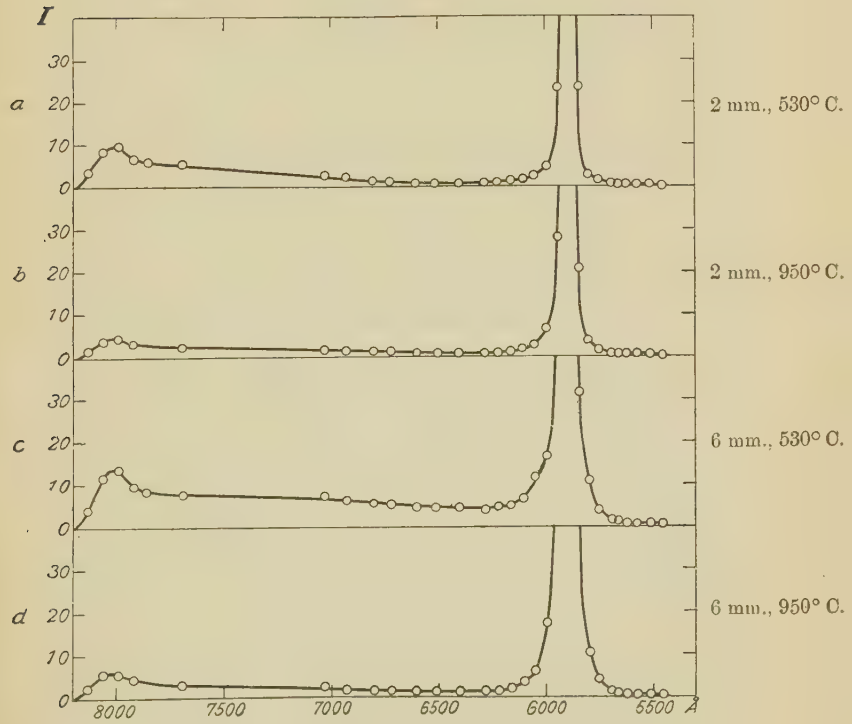
(IV.) Next, on investigating the effect of the exciting current density, the following results were obtained. The effect of current is not so great in the range between 20 to 100 milliamps., and the increase of current gives rise to intensifications of the continuous as well as of the discrete bands as a whole, but the former is relatively enhanced in comparison with the latter. Increasing the current also gives rise to a diffusion of the insensity maxima in the resonance series B, together with a relative reduction of it compared with the band system A. In other words, the band spectrum due to the sodium molecules directly excited by the impacts of electrons appears enhanced in the case of a high current density.

(V.) The intensity distribution in each kind of spectrum. The results of the photometric measurement of the spectrograms are shown in fig. 4, in which the small fluctuations, due to the band-heads or the maxima in the discrete bands A and B, have been straightened out. It is also to be remarked that, as the intensities in the regions very near the D-lines are very great, and also, as the self-absorption taking place in addition, it is difficult to determine accurately the intensity in these regions from the photometric measurement. However, it is certain that the intensity at the wave-length between the D_1 and D_2 lines ($\sim$5893 Å.) is always greater than 10^3 in the order of magnitude, in the arbitrary unit shown in the figure. The intensity of the core of D-line emitted by the elementary act may be markedly different in the four cases of the different experimental conditions described in the figure, but it may be safely said that its value is always far greater than this value.

The above four curves represent the intensity distributions of the total spectrum directly obtained from the photographic plates, including discrete bands A and B and the continuous spectrum C. The comparison of the intensities of certain wave-length in *a*, *b*, *c*, and *d* with each other also gives, approximately, a correct result. It is difficult to find the intensity distribution of each of the three kinds of spectrum

from this figure directly, for they are superimposed on each other, especially in the case of a lower temperature. But as the dispersion of the spectrograph used was fairly great, it is possible to discriminate between the discrete bands and the continuous spectrum and to ascertain, in a qualitative sense, the intensity distribution of each class of spectrum separately.

Fig. 4.



As described above, the three kinds of spectrum due to the diatomic molecules of sodium are generally emitted intensely under high vapour pressure. On account of the fact that the discrete bands A and B are very faintly emitted at a high temperature, the state of excitation of the vapour, which is at a very high temperature and comparatively high pressure, is the most favourable condition for the observation of the intensity distribution of the continuous spectrum. It was observed that this intensity distribution is nearly symmetrical on both sides of the wave-lengths and is at a

remarkable maximum at the D-lines, and the rise of temperature, without altering the pressure or the increase of pressure, the temperature being kept constant, causes the relative or absolute enhancement, and, therefore, the (apparent) broadening of the continuous spectrum. In the region of the shorter wave lengths, its intensity decreases more or less abruptly at about 5600 Å., and fades away with a comparatively sharp edge at about 5470 Å., and there is a flat and hazy maximum at about 5500 Å., while, on the side of the longer wave-lengths, its intensity is gradually and uniformly decreased, and it can be traced up to the region of about 7200 Å. The determination of the longer wave-length limit of this continuous spectrum is very difficult, because its relative intensity to the discrete band of the same wave-length becomes smaller and smaller as we proceed towards the longer wave-lengths, but it is certain that the spectrum does not extend beyond the region at about 8200 Å, as any plate when over-exposed shows no trace of the continuous or discontinuous spectrum beyond this region.

It is difficult to observe the rotation structure of the discrete bands on the plate taken with the prism spectrograph, though it is not difficult to distinguish between the continuous spectrum and the discrete bands. Therefore the second-order spectrogram by the concave grating was photographed, and there were found many lines crowded together at the positions corresponding to the intensity maxima of the resonance series B, while the weak lines are distributed in the intervening space between these maxima. The results of observation are in perfect accordance with the result given above, and it is thus observed that this rotation structure is always sharply defined without showing any appreciable broadening of the band-lines as the result of the increase of pressure, and also of the rise of the temperature, until the condition (very high temperature) is arrived at where the discrete bands and therefore each separate band-line ultimately disappears. From the results above described, it is proved that there is no direct connexion between the structure of the discrete bands and the continuous spectrum.

It is remarked here that, when the vapour pressure is increased gradually from a very small value, each atomic D-line is broadened out unsymmetrically, and the line D_1 (5896 Å.) is broadened out more intensely towards the red, while the line D_2 (5890 Å.) is broadened out more intensely towards the violet. The observed distribution of intensity of the spectrum, which is the superposition of the continuous

spectra about both resonance lines, has already been described.

(V.) Green-band system and the accompanying continuous spectrum. In the wave-length region, where the green-band system mainly exists, there was found a continuous spectrum and the discrete bands, and the intensity of the former gradually increases, while that of the latter gradually fades away, towards the D-lines In the fluorescence as well as in the absorption, both the green- and the red-band systems appear with nearly the same intensity, or rather the former is more intensely excited than the latter, while, under the conditions of this experiment, the green-band system and the continuous spectrum of its background are very faintly excited, in comparison with the red-band system and the accompanying continuous spectrum. The changes of intensity of the discrete band system and of the continuous spectrum due to the variation of the temperature, the vapour pressure, and the density of the exciting current are exactly similar to those in the case of the red-band system and the continuous spectrum which forms its background. It was possible to trace the continuous spectrum up to the region at about 4550 Å., where its intensity fades away, but this shorter wave-length limit is not so sharp. As there is an intense continuous spectrum with the violet limit at about 5500 Å., as above described, it is impossible to observe the intensity distribution of this continuous spectrum in the longer wave-length region beyond the position at 5500 Å., but it seems that this continuous spectrum starts from the region very near the D-lines, and that its intensity gradually decreases towards the violet.

Discussion of the Results.

Obviously, for the interpretation of the results above described, it will be necessary to consider only the molecular terms, which are derived from a sodium atom in the normal ^2S state, and another sodium atom in the normal ^2S state, or in the lowest ^2P state. By applying the so-called Hund-Wigner-Witmer rule, two molecular terms, $^1\Sigma_g^+$ and $^3\Sigma_u^+$, are derived from the two normal ^2S atoms, and eight molecular terms, $^1\Sigma_g^+$, $^1\Sigma_u^+$, $^1\Pi_g$, $^1\Pi_u$, $^3\Sigma_g^+$, $^3\Sigma_u^+$, $^3\Pi_g$, and $^3\Pi_u$, are derived from the atomic combination $^2S + {}^2P$ *. One part of these molecular terms is stable, and each of their potential energy

* R. S. Mulliken, Phys. Rev. xxxvi. p. 1440 (1930).

curves has a deep or shallow minimum, while the remaining ones are unstable continuous terms, and each of their potential energy curves is repulsive, and has no conspicuous minimum. Taking into consideration the relative positions and the forms of the potential curves of the molecular states above described, as well as the concentrations of the excited molecules in the initial states, and also the transition probability, the observed behaviour of the discrete bands and of the continuous spectrum of sodium may be explained.

As the lower molecular levels of the transitions are only the $^1\Sigma_g^+$ and the $^3\Sigma_u^+$ states, only the four transitions $^1\Sigma_u^+ \rightarrow {}^1\Sigma_g^+$, $^1\Pi_u \rightarrow {}^1\Sigma_g^+$, $^3\Sigma_g^+ \rightarrow {}^3\Sigma_u^+$, and $^3\Pi_g \rightarrow {}^3\Sigma_u^+$, among all the combinations between the molecular terms dissociating into the atomic combinations $^2S+{}^2S$ and $^2S+{}^2P$, are permissible or favoured on the selection principles. As two of the four transitions above mentioned are well known as the red-band system ($^1\Sigma_u^+ \longleftrightarrow {}^1\Sigma_g^+$) and the green-band system ($^1\Pi_u \longleftrightarrow {}^1\Sigma_g^+$), and the vibrational and rotational analyses have been satisfactorily completed, it is possible to plot the potential energy curves of the molecular states $^1\Pi_u$, $^1\Sigma_u^+$, and $^1\Sigma_g^+$, as shown in the paper of Loomis and Nile *, or as shown in fig. 5, in which the energies near the region of the equilibrium positions of the nuclei were calculated by the ordinary method with the data recently obtained by Loomis † and others, while the other parts of the potential curves were calculated by Morse's formula. By these curves it is clearly shown that there are sharp edges at about 5500 and 8150 Å. respectively, in the red-band system. This fact has been confirmed by Loomis and Nile for the first time, and well explained by them from the relative positions and the forms of the potential curves. As the lower $^3\Sigma_u^+$ is an unstable continuous term ‡, the transitions $^3\Sigma_g^+ \rightarrow {}^3\Sigma_u^+$ and $^3\Pi_g \rightarrow {}^3\Sigma_u^+$ from the upper stable or unstable state give no discrete band system, and may give rise to a continuous spectrum, but the continuous spectrum described in this experiment is not explained by these transitions only, because this spectrum has a sharp edge at about 5500 Å., and also extends far into the region of the shorter wavelengths (4550 Å.).

There is some concentration of the normal molecules in the sodium vapour, and the stable excited molecules in the

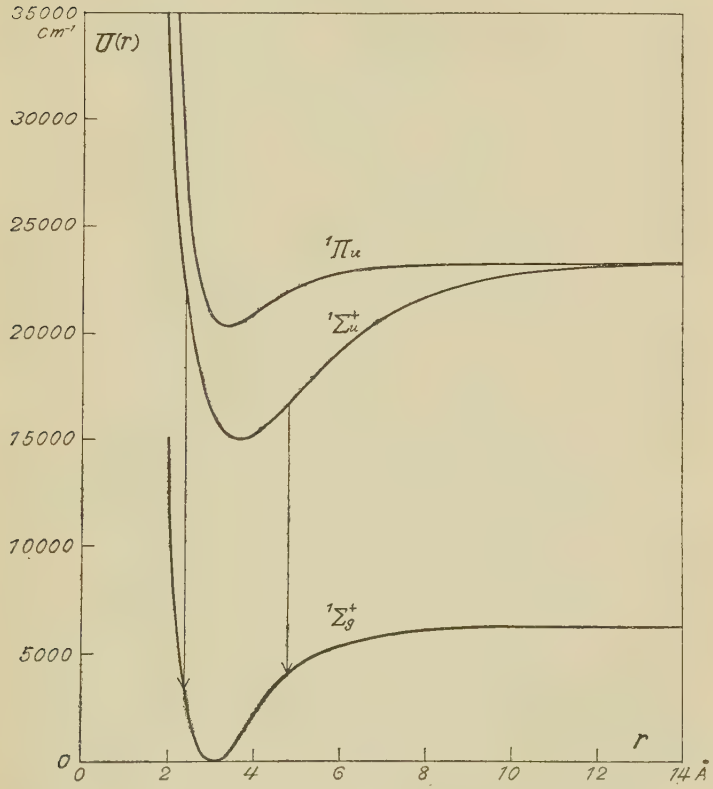
* F. W. Loomis & S. W. Nile, Phys. Rev. xxxii. p. 873 (1928).

† F. W. Loomis, Phys. Rev. xxxviii. p. 2153 (1931).

‡ G. Herzberg, *Zeits. f. Phys.* lvii. p. 601 (1929) F. Hund, *Zeits. f. Phys.* lxiii. p. 719 (1930).

states $^1\Pi_u$ and $^1\Sigma^+_u$ are produced by the collisions of normal molecules with electrons or by the triple collisions. Further, excited molecules in the $^1\Sigma^+_u$ state with very limited ranges of values of rotational and vibrational quantum numbers are produced from the stable molecules in consequence

Fig. 5.



of collisions of the second kind with the excited atoms in the ^{2}P state, or else of the absorption of the energy of the broadened D-lines, and the behaviour of the discrete bands belonging to the red- and the green-band systems, can be well explained, as described in the researches of Loomis and others, aud also in this report.

In addition to the emission of the discrete bands by excited stable molecules in the $^1\Pi_u$ and $^1\Sigma^+_u$ states, there

is also possible the emission of the continuous spectrum by the excited ^2P atoms during their collisions with other normal atoms, viz., by the excited quasi-molecules * in the $^1\Sigma_u^+$, $^1\Pi_u$, $^3\Sigma_g^+$, and $^3\Pi_g$, states. Which term will be given, *i. e.*, in other words, along which potential curve the two atoms approach each other, when the two atoms, each with a specified electron configuration and in a definite state—characterized by the quantum numbers L, S, and J—are in collision, depends on the mutual (quantized) orientations of the orbital angular momentum vectors and of the spin angular momentum vectors of the electrons of the two atoms. The continuous spectrum with a maximum at the D-lines may certainly be explained by the emission of the excited quasi-molecules above mentioned. As some of the potential curves descend, while the others ascend, when the two atoms (^2S and ^2P) approach each other closely, while the probability of electronic transitions is greater when they are far apart, the reason why the continuous spectrum has an intense maximum at the atomic lines, and also why it broadens towards the wave-lengths on both sides of the lines, when the temperature or the pressure of the sodium vapour is raised, is easily understood.

But it is problematical whether all the potential curves have an equal significance, and the concentration of the quasi-molecules in each excited state separately, as well as the transition probability, must be taken into consideration. In the case of the diatomic molecules of mercury, cadmium, zinc, and magnesium, the intensity distribution of the continuous spectrum about the resonance line is always unsymmetrical with fairly sharp edge on the side of the shorter wave-lengths; but if we observe the neighbouring regions of the resonance line and compare the intensities of the wave-numbers $\nu+\Delta\nu$ and $\nu-\Delta\nu$, where ν is that of the resonance line, the continuous spectrum about the resonance line $1^1S_0-2^1P_1$ always broadens out nearly symmetrically in all the metals, but the resonance line $1^1S_0-2^3P_1$ broadens out extremely unsymmetrically in mercury, to a less degree in cadmium, while in the case of zinc it broadens out nearly symmetrically. If, judging not only from the attractive potential curve, but also from the repulsive potential curve, we consider that electronic transitions occur,

* M. Born and J. Franck, *Zs. f. Phys.* xxxi. p. 411 (1925). A. Jablonski, *Zs. f. Phys.* lxx. p. 723 (1931); W. Weizel, Phys. Rev. xxxviii. p. 642 (1931). H. Kuhn and O. Oldenberg, Phys. Rev. xli. p. 72 (1932). H. Kuhn, *Zs. f. Phys.* lxxvi. p. 782 (1932).

as Kuhn * has already suggested, then the unsymmetrical broadening of the continuous spectrum may well be understood, because the degree of symmetry or asymmetry can be explained by the forms and relative positions of the potential curves, taking into consideration the (instantaneous) concentrations of the quasi-molecules with specified configurations and with specified values of inter-nuclear distance, and also the transition probability.

From the experimental result that, in the case where the temperature of the vapour is very high and all the discrete bands almost completely disappear, the intense continuous spectrum about the D-lines terminates with a sharp edge just at ~5500 Å., viz., at exactly the same wave-length where the discrete red-band system suddenly terminates, it seems that the emission of the main part of the strong continuous spectrum in the background of the red-band system is closely connected with the electronic transitions $^1\Sigma_u^+ \rightarrow {}^1\Sigma_g^+$ of the quasi-molecules. The reason why the sharp edge is lacking on the side of the longer wave-lengths of this continuous spectrum, while there is a strong sharp edge on the side of the shorter wave-lengths, will follow from the Franck-Condon principle, as the emission of the continuous spectrum in the region near the infra-red edge is due to the transitions from the positions of the potential curve, where the two atoms approach or depart with considerable relative velocity, while the emission of the continuous spectrum near the edge on the side of the shorter wave-lengths may perhaps be due to the electronic transitions from the positions of the potential curve, where the two colliding atoms have a smaller relative velocity, on account of these positions being, perhaps, very near the horizontal asymptote.

In the case of a weak continuous spectrum, which is developed in the background of the green-band system, as its intensity gradually decreases towards the side of the shorter wave-lengths and terminates just at ~4550 Å., where the discrete green-band system fades away, it is natural to consider that the continuous spectrum in the background of the green-band system has an intimate connection with the latter band system, and that it is emitted by the transitions $^1\Pi_u \rightarrow {}^1\Sigma_g^+$ of the quasi-molecules.

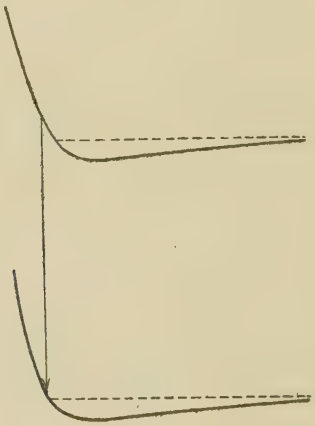
Besides the continuous spectrum due to the transitions $^1\Sigma_u^+ \rightarrow {}^1\Sigma_g^+$ and $^1\Pi_u \rightarrow {}^1\Sigma_g^+$, there is also possible the emission of a continuous spectrum on the side of the longer wave-

* H. Kuhn, *Zs. f. Phys.* lxxii. p. 462 (1931).

lengths, or in the vicinity of the D-lines, in consequence of the transitions ${}^3\Sigma_g^+ \rightarrow {}^3\Sigma_u^+$ and ${}^3\Pi_g \rightarrow {}^3\Sigma_u^+$ of the quasi-molecules, and the experimental fact that, as we proceed towards the side of the shorter wave-lengths from the D-lines, the intensity of the continuous spectrum decreases more or less abruptly at about 5600 Å., may be explained by the electronic transitions from one of the upper states (${}^3\Pi_g$?), the potential curve of which and that of the lower ${}^3\Sigma_u^+$ state are assumed, in the shapes and in the relative positions, as shown in fig. 6.

From the result of this experiment, it is probable that, in collision, the normal ${}^2S$ atom and the excited atom in the lowest ${}^2P$ state approach each other closely with every possible orientation of the orbital angular momentum vectors

Fig. 6.



and of the spin vectors; but, of all these, those mutual orientations, from which the stable excited ${}^1\Sigma_u^+$ state is derived when the vibration and rotation are quantized seem to be the ones most likely to occur. It is very probable that the stable excited molecules in the ${}^1\Sigma_u^+$ state, which resulted in consequence of the triple collisions, or, in other words, in consequence of the quantization of quasi-molecules, play an important part in the emission of the discrete red-band system in this experiment.

The main part of the continuous spectrum emitted from the sodium arc between carbon electrodes, when the density of sodium vapour is considerably high, is evidently the superposition of the spectrum due to the temperature radiation

and that due to the quasi-molecules in the states $^1\Sigma_u^+$, $^1\Pi_u$, etc., and, if we remove the continuous spectrum due to the former from the observed intensity distribution, it is known that there is a very intense maximum at the D-lines and a sharp edge at about 5500 Å., and the intensity is gradually decreased towards the side of the longer wave-lengths in the spectral regions where the red-band system exists, and also towards the violet side in the regions where the green-band system exists ; just as in the above experiment. It was also noted that, in the case of the sodium flame, the continuous spectrum about the D-lines has also the similar distribution of intensity as that obtained in this experiment. Leder has already observed the flat maximum of intensity at about 5500 Å. in his experiment on the flame spectra, and he considered that it is a halo of the lines 5533–27 Å. However, as these lines are due to P–P combinations, they should appear very weakly, and it will be reasonable to consider that the flat maximum at ~5500 Å. is that of the red continuous spectrum ($^1\Sigma_u^+ \rightarrow {}^1\Sigma_g^+$) of quasi-molecules, and is not due to the broadening of the atomic lines 5533–27.

Evidently, in the sodium arc, or in the sodium flame, in air, collisions between the excited sodium atoms in the ^{2}P state and the atoms or molecules of other elements may also be taking place, and the continuous spectrum in the neighbourhood of the D-lines may be emitted during these collisions, but it is certain, from the observed result of the intensity distribution, that the quasi-molecules of sodium make the great contribution to the emission of a continuous spectrum. As the sodium vapour is under a very high temperature in the arc or in the flame, the reason why the discrete bands are not emitted in it, as the results of observations show, is easily understood from the result of this experiment.

Exactly similar results are obtained in potassium. It was observed that a continuous spectrum is emitted in the region from 7590 to 7900 Å. with an intense maximum at the resonance lines. It may also be probable that the quasi-molecules in the $^1\Sigma_u^+$ state exist in great concentration. From the forms and relative positions of the two potential curves in the states $^1\Sigma_u^+$ and $^1\Sigma_g^+$, it may follow from the experimental results that the discrete band system $^1\Sigma_u^+ \rightarrow {}^1\Sigma_g^+$ is emitted on the side of the longer wave-length of the resonance lines with a sharp edge in the infra-red region, while, contary to the result in the case of sodium, no sharp edge is observed far apart from the resonance lines on the side of

the shorter wave-lengths *. Therefore, the continuous spectrum due to the transitions ${}^1\Sigma_u^+ \rightarrow {}^1\Sigma_g^+$ of quasi-molecules has also no sharp edge on the side of the shorter wave-lengths of the resonance lines, and also, similarly to the case of sodium, it will have no sharp edge on the side of the longer wave-lengths.

In conclusion, the writer wishes to express his hearty thanks to Prof. J. Okubo for his help and advice during the course of the experiment.

Sendai, Japan.
September 1932.

XLIX. *Uncertainty Relations and Volume of Photons.* *By* D. Meksyn, *Ph.D., D.Sc.* †

§ 1. *Summary.*

THE aim of this paper is to apply the laws of quantisation of an electromagnetic field in order to derive certain properties of photons.

Exchange relations between the electric and magnetic field of a photon, associated with a plane wave, are obtained, and their bearing to Jordan's, Pauli's, and Heisenberg's ‡ exchange relations is discussed.

In our case the uncertainty between the electric and magnetic force of a *single* photon is proportional to its energy; whence follows that the division of the field of a photon in two separate electric and magnetic fields has no meaning if applied to a single photon.

It is further shown that the exchange relations for the electromagnetic field due to the superposition of n similar photons, each of energy $h\nu$, is proportional to $\frac{1}{n}$; or, if n is sufficiently great, the separation of the electromagnetic field in an electric and magnetic field can be carried out with a sufficient degree of approximation.

The volume of a photon is evaluated, and is found to be equal to $\frac{3\lambda^3}{8\pi}$, where λ is its wave-length.

* F. W. Loomis and R. E. Nusbaum, Phys. Rev. xxxix. p. 89 (1932).

† Communicated by the Author.

‡ P. Jordan and W. Pauli, *Zs. f. Phys.* xlvii. p. 151 (1928); W. Heisenberg and W. Pauli, *ibid.* lvi. p. 1 (1929); lix. p. 168 (1930).

§ 2. *Exchange Relations for Photons.*

Consider a plane standing electromagnetic wave. Its vector potential is of the form

$$\mathrm{U} = \mathrm{A}u(t)\sin\left[\frac{2\pi\nu}{c}(\alpha, \mathrm{X}) + \beta\right], \quad . \quad . \quad . \quad (1)$$

where X is a vector with components x, y, z, α is a unit vector giving the direction of the standing wave, A is a unit vector giving the direction of vibration of the electric force, and it is normal to α. The factor $u(t)$ which gives the dependence of U on time is generally a sine function of t.

The general expression for the vector potential is

$$\left.\begin{aligned} \mathrm{U} &= \Sigma_s \mathrm{A}_s u_s(t)\sin\theta_s, \\ \theta_s &= \frac{2\pi\nu_s}{c}(\alpha_s, \mathrm{X}) + \beta_s, \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (2)$$

the electromagnetic field is given by

$$\mathrm{E} = -\frac{1}{c}\frac{\partial \mathrm{U}}{\partial t}; \qquad \mathrm{H} = \mathrm{rot}\,\mathrm{U}; \quad . \quad . \quad . \quad (3)$$

hence from (2) and (3) we obtain

$$\left.\begin{aligned} \mathrm{E} &= -\Sigma_s \frac{1}{c}\mathrm{A}_s \dot{u}_s \sin\theta_s, \\ \mathrm{H} &= \Sigma_s \frac{2\pi\nu_s}{c}[\alpha_s, \mathrm{A}_s] u_s \cos\theta_s. \end{aligned}\right\} \quad . \quad . \quad . \quad (4)$$

If the field is enclosed in a space of volume V its average energy is

$$\mathrm{W} = \mathrm{V}.\frac{\overline{\mathrm{E}^2} + \overline{\mathrm{H}^2}}{8\pi},$$

or, remembering that

$$\overline{\sin^2\theta_s} = \overline{\cos^2\theta_s} = \tfrac{1}{2}; \quad [\alpha_s, \mathrm{A}_s]^2 = 1,$$

we obtain

$$\mathrm{W} = \frac{\mathrm{V}}{8\pi c^2}\Sigma_s\left(\tfrac{1}{2}\dot{u}_s^2 + 2\pi^2\nu_s^2 u_s^2\right); \quad . \quad . \quad . \quad (5)$$

this is equivalent to the energy of a system of harmonic oscillators of frequency ν_s. Imposing now the ordinary exchange relations on the coordinates and momenta of these

equivalent oscillators, and constructing a combined Hamiltonian function for matter and radiation, one can derive the laws of interaction between matter and radiation †. The above-sketched calculation is well known, and is given here for the sake of convenience.

We have to modify now the equations (1–5) for the purpose of deriving properties of photons.

We consider that in our space there are photons only of a certain frequency ν, and, following the same procedure as the quantum dynamics of a particle, we assume that the wave function associated with a photon is of the form

$$\sim \exp \frac{2\pi\nu i}{c}[(\alpha, \mathrm{X}) + \beta - ct],$$

where α, X, and β have the same meaning as before ‡. We need not show explicitly the dependence of the wave function upon time, and we put, accordingly,

$$\left.\begin{aligned} \mathrm{U} &= \mathrm{A}u(t)e^{i\theta}, \\ \theta &= \frac{2\pi\nu}{c}[(\alpha, \mathrm{X}) + \beta], \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (2a)$$

whence

$$\left.\begin{aligned} \mathrm{E} &= -\frac{\mathrm{A}}{c}\dot{u}e^{i\theta}, \\ \mathrm{H} &= \frac{2\pi\nu}{c} i[\alpha, \mathrm{A}]e^{i\theta}u. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (4a)$$

The energy of the photons is given in this case by

$$\mathrm{W} = \mathrm{V}.\frac{\mathrm{E}.\mathrm{E}^* + \mathrm{H}.\mathrm{H}^*}{8\pi},$$

where E^*, H^* are conjugate to E and H. Following the same procedure that led us to (5), we obtain for the energy W, instead of (5),

$$\mathrm{W} = \frac{\mathrm{V}}{4\pi c^2}(\tfrac{1}{2}\dot{u}^2 + 2\pi^2\nu^2u^2), \quad . \quad . \quad . \quad . \quad (6)$$

which is double of (5).

Put in the ordinary way

$$v = \frac{\partial \mathrm{W}}{\partial \dot{u}} = \frac{\mathrm{V}}{4\pi c^2}\dot{u}, \quad . \quad . \quad . \quad . \quad . \quad (7)$$

† An account of this theory is given by Fermi, "Quantum Theory of Radiation," 'Reviews of Modern Physics,' no. 1, p. 87 (1932).

‡ C. G. Darwin, Proc. Roy. Soc. A, cxxxvi. p. 36 (1932).

and denote

$$u = \left(\frac{4\pi c^2}{\mathrm{V}}\right)^{\frac{1}{2}} q \; ; \quad v = \left(\frac{\mathrm{V}}{4\pi c^2}\right)^{\frac{1}{2}} p. \quad . \quad . \quad . \quad (8)$$

From (4*a*), (6), (7), and (8) we obtain the following expressions for the energy and the electromagnetic force of our wave :

$$\mathrm{W} = \tfrac{1}{2}p^2 + 2\pi^2\nu^2 q^2, \quad . \quad . \quad . \quad . \quad . \quad (9)$$

$$\mathrm{E} = -\frac{\mathrm{A}}{c}\left(\frac{4\pi c^2}{\mathrm{V}}\right)^{\frac{1}{2}} p, \quad \mathrm{H} = \frac{2\pi\nu}{c}[\alpha, \mathrm{A}]\left(\frac{4\pi c^2}{\mathrm{V}}\right)^{\frac{1}{2}} q. \quad (10)$$

We impose on p and q the ordinary exchange relations

$$qp - pq = \frac{ih}{2\pi}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

From (9), (10), and (11) we find

$$\mathrm{EH} - \mathrm{HE} = \frac{4\pi i h\nu}{\mathrm{V}} \mathrm{A}.[\alpha, \mathrm{A}].$$

Now let α be in the direction of x and A in the direction of y ; they are unit vectors. If we are dealing with a single light quantum, or with many similar light quanta, we obtain for the above choice of directions at α and A,

$$\mathrm{A}.[\alpha, \mathrm{A}] = 1,$$

and hence the final result

$$\mathrm{EH} - \mathrm{HE} = \frac{4\pi i h\nu}{\mathrm{V}}. \quad . \quad . \quad . \quad . \quad . \quad (12)$$

E and H are taken along their directions, or (12) is an intrinsic exchange relation.

§ 3. *Interpretation of the Exchange Relation.*

In order to make clear the physical meaning of the uncertainty relations which follow from (12) we take as a unit of the electric and magnetic force their values according to the classical electrodynamics.

For a plane wave moving along the axis of x the amplitudes of the electric and magnetic forces are equal, or

$$\mathrm{Y}_0 = \gamma_0 ;$$

the electromagnetic force is accordingly

$$\mathrm{Y}_0 \sin\frac{2\pi\nu}{c}(x - ct) \; ; \quad \gamma_0 \sin\frac{2\pi\nu}{c}(x - ct),$$

and the average density of energy is

$$\frac{Y_0^2}{8\pi} = \frac{\gamma_0^2}{8\pi};$$

or the total energy of a volume V is

$$W = V.\frac{Y_0^2}{8\pi};$$

whence

$$Y_0 = \gamma_0 = \sqrt{\frac{8\pi W}{V}}. \quad . \quad . \quad . \quad . \quad (13)$$

Thus (13) gives us the classical values of the amplitudes of the electromagnetic field of frequency ν included in a volume V and having a total energy W.

The uncertainty relation which follows from (12) is

$$\Delta E.\Delta H \eqsim \frac{8\pi^2 h\nu}{V}. \quad . \quad . \quad . \quad . \quad . \quad (14)$$

As it is more important to know the relative uncertainties of measurement we introduce the quantities

$$\frac{\Delta E}{Y_0} = \Delta E_0, \qquad \frac{\Delta H}{\gamma_0} = \Delta H_0, \quad . \quad . \quad . \quad (15)$$

where ΔE_0 and ΔH_0 are the ratios between the uncertainties of the electric and magnetic fields and their average values as given by the classical electrodynamics.

From (13), (14), and (15) we obtain

$$\Delta E_0.\Delta H_0 \eqsim \frac{\pi h\nu}{W}. \quad . \quad . \quad . \quad . \quad . \quad (16)$$

We consider the meaning of (16) in the case of a single light quantum and in the case of an electromagnetic field produced by any number of similar light quanta.

a. *A single Photon.*

The energy $W = h\nu$, and (16) becomes

$$\Delta E_0.\Delta H_0 \eqsim \pi, \quad . \quad . \quad . \quad . \quad . \quad (17)$$

or the best exactness of measurement which can be obtained both for the electric and magnetic forces of a single photon is about 180 per cent. This shows that the division of an electromagnetic field in separate electric and magnetic fields has no meaning for a single photon.

Several papers appeared recently which sought to set up a wave equation for photons analogous to Dirac's wave equation for matter †.

It might appear at first that there is no analogy between these two cases. In Schrödinger's equation we can measure only $\psi\psi^*$, but not separately ψ and ψ^*, while, as in the case of an electromagnetic field, we can it seems measure not only $\frac{E^2+H^2}{8\pi}$, but separately E and H. Attention to this point was drawn recently by Ehrenfest ‡.

If our uncertainty relation is correct it clears up the above-mentioned difficulty; in both cases only a definite combination of the field functions can be measured.

b. *Many Photons.*

Let us consider the case of n similar photons of frequency ν, enclosed in a volume V. Their total energy $W = nh\nu$, and (16) becomes

$$\Delta E_0 . \Delta H_0 \geqq \frac{\pi}{n}, \quad . \quad . \quad . \quad . \quad . \quad (18)$$

or, when n tends to infinity, the electric and magnetic forces commute; for large n the electric and magnetic fields can be measured with a sufficient degree of approximation.

§ 4. *Comparison with Heisenberg's Uncertainty Relations.*

Heisenberg's uncertainty relation is

$$\Delta E_x . \Delta H_y \geqq \frac{hc}{(\delta l)^4}, \quad . \quad . \quad . \quad . \quad . \quad (19)$$

where δl is an element of length §. It appears somewhat different from our equation (14). We shall show, however, that our exchange relation (12) for a photon can be derived from the equation (W. Heisenberg, *ibid.* p. 51)

$$E_i(r)\Phi_a(s) - \Phi_a(s)E_i(r) = -2hci\delta_{rs}\delta_{ai}\frac{1}{\delta v}, \quad . \quad (19a)$$

† Landau and Peierls, *Zs. f. Phys.* lxii. p. 188 (1930); Oppenheimer, Phys. Rev. xxxviii. p. 725 (1931); Darwin, Proc. Roy. Soc. A, cxxxvi. p. 36 (1932).

‡ P. Ehrenfest, *Zs. f. Phys.* lxxviii. p. 558 (1932).

§ W. Heisenberg, 'The Physical Principles of the Quantum Theory,' p. 50.

where E and Φ are the electric force and the vector potential, δv an element of volume of the field, and

$$\delta_{rs} = \begin{cases} 1 & \text{for} \quad r = s, \\ 0 & \text{for} \quad r \neq s. \end{cases}$$

We insert in (19 *a*) our values for E and Φ. From (2 *a*) and (4 *a*) we have

$$\Phi_y = u(t)e^{i\vartheta} \; ; \quad \mathrm{E}_y = -\frac{\dot{u}}{c}e^{i\vartheta} \; ; \quad \mathrm{H}_z = \frac{2\pi\nu i}{c}ue^{i\vartheta}. \quad (19\,b)$$

From (19 *b*) we see that

$$\Phi_y = \frac{c}{2\pi\nu i}\mathrm{H}_z. \quad . \quad . \quad . \quad . \quad . \quad (19\,c)$$

In equation (19 *a*) E and Φ stand for the amplitudes of the electric force and the vector potential; combining (19 *a*), (19 *b*), and (19 *c*), we obtain (dropping the phase factors)

$$\mathrm{E}_y\mathrm{H}_z - \mathrm{H}_z\mathrm{E}_y = -\frac{4\pi h\nu i}{\delta v},$$

which is our result (12).

We can see now why our uncertainty relations (17) and (18) differ from (19).

Heisenberg considers the uncertainty of measurement for an arbitrary electromagnetic field, without specifying as to how the field is produced; the relation (19) gives therefore the best exactitude of measurement which can be obtained.

In our case the field is produced either by one or by a certain number n of photons; this enables us to obtain additional information concerning photons.

§ 5. *Quantum Equations for Photons.*

We recapitulate the exchange relations between the electric and magnetic forces of photons.

In our case the photon is a plane wave of frequency ν; it moves along the axis of z. The electromagnetic field is of the form

$$\mathrm{X} = \mathrm{A}e^{\frac{2\pi i\nu}{c}[(z-ct)+\beta]} \quad . \quad . \quad . \quad . \quad . \quad (20)$$

The amplitudes satisfy the following exchange relations:—

$$\left.\begin{aligned} \mathrm{E}_x\mathrm{H}_y - \mathrm{H}_y\mathrm{E}_x = \frac{4\pi ih\nu}{\Omega}; \quad \mathrm{H}_x\mathrm{E}_y - \mathrm{E}_y\mathrm{H}_x = \frac{4\pi ih\nu}{\Omega}; \\ \mathrm{E}_x\mathrm{H}_x - \mathrm{H}_x\mathrm{E}_x = 0 \; ; \quad \mathrm{E}_y\mathrm{H}_y - \mathrm{H}_y\mathrm{E}_y = 0 \; ; \\ \mathrm{E}_x\mathrm{E}_y - \mathrm{E}_y\mathrm{E}_x = 0 \; ; \quad \mathrm{H}_x\mathrm{H}_y - \mathrm{H}_y\mathrm{H}_x = 0, \end{aligned}\right\} \quad (21)$$

where Ω is the volume of a photon.

In order to find the quantized values of the electromagnetic field E and H we can make use of (10) and (11); it is possible, however, to find them in a more direct way. To this end it is only necessary to add to the equation (21) the expression for the energy of a photon,

$$W = \frac{\Omega}{8\pi}(E_x^2 + E_y^2 + H_x^2 + H_y^2), \quad . \quad . \quad . \quad (22)$$

where, of course, the E's and H's are here matrices.

The solution of (21) and (22) gives us the required values for the E's and H's and for the energy of a photon.

§ 6. *Volume of Photons.*

We can make use of the values of E obtained from (21) and (22) in order to derive the laws of emission and absorption of radiation. It will lead us to the well-known results; but as the volume of photons appears in our equations (21) and (22), we could expect to find the value of this volume from the laws of interaction between matter and radiation. This appears to be the case.

We can also obtain the volume of a photon in the following simple way.

Let us have radiation enclosed in a certain volume V. The number of vibrations between the frequencies ν and $\nu + d\nu$ is equal to

$$\frac{8\pi\nu^2 d\nu}{c^3} \quad . \quad . \quad . \quad . \quad . \quad . \quad (23)$$

for a unit of volume; or for a wave-length λ the corresponding number of vibrations between λ and $\lambda + d\lambda$ is

$$\frac{-8\pi d\lambda}{\lambda^4}.$$

Let the volume of a photon of wave-length λ be Ω; then the maximum number of possible vibrations of wave-length λ in a unit of volume is $\frac{1}{\Omega}$. Suppose now that the volume of a photon is equal to $\Omega = a\lambda^3$, where a is a constant, evidently

$$\frac{d}{d\lambda}\left(\frac{1}{a\lambda^3}\right) = -\frac{8\pi d\lambda}{\lambda^4},$$

or $a = \frac{3}{8\pi}$; and hence the required volume is

$$\Omega = \frac{3\lambda^3}{8\pi}.$$

In this derivation of the volume of a photon we have assumed that no two photons can be at the same place, and that in the case when the maximum number of photons is present in space they completely occupy this space, *i. e.*, there are no intervals between photons.

We know that only a finite number of vibrations or photons can be present in a unit of volume; the above assumptions seem to be the simplest which could account for the existence of such a number from the particle point of view.

However natural these assumptions are they are not necessary in order to obtain the volume of a photon.

These assumptions and the volume of photons can be derived from the laws of emission of radiation.

If we evaluate the probabilities of emission of photons by a method similar to that given, for instance, in Weyl's work †, but making use of (21) and (22), we find an expression of the form $\sum_s \frac{h\nu_s}{\Omega_s}$, where ν_s lies between ν and $\nu + d\nu$, and Ω_s is the corresponding volume of a photon of frequency ν_s. In order to arrive at the correct expression for the probability of emission we must put

$$\sum_s \frac{1}{\Omega_s} = \frac{8\pi^2 \nu d\nu}{c^3},$$

whence follow the above assumptions and the volume of a photon.

As the computation is rather lengthy and otherwise identical with Weyl's derivation we do not reproduce it here.

Attempts have been made to find the size of a photon. E. O. Lawrence and J. W. Beams ‡ sent a ray of light through a shutter consisting of two crossed Nicol prisms and two Kerr cells placed between them at right angles to each other; the shutter was capable of interrupting a beam of light more than 10^9 times per second. Lawrence and Beams found that light quanta passed through the cells in less than

† H. Weyl, 'The Theory of Groups and Quantum Mechanics,' 1930, § 13.

‡ Phys. Rev. xxxii. p. 478 (1928).

10^{-9} seconds, so that the quanta are not more than a few centimetres in length.

Similar experiments were made by Breit †. The results of these experiments are not against our value for the volume of a photon, but they do not confirm it. The exactitude of the measurements was not sufficient in order to check our result.

If our expression for the volume of a photon is correct there appears to be a certain resemblance between the structure of photons and electrons.

We usually consider that the size of a proton is much smaller than the size of an electron, although its mass (energy) is much greater. A similar relation exists between the size and energy of a photon; the greater its energy the smaller its volume.

I wish to express my thanks to Prof. Darwin for his invaluable criticism of this paper and for many helpful suggestions.

London.

L. *Probe Measurements in the Sodium Arc.* *By* F. H. Newman, *D.Sc., F.Inst.P., Professor of Physics, University College, Exeter* ‡.

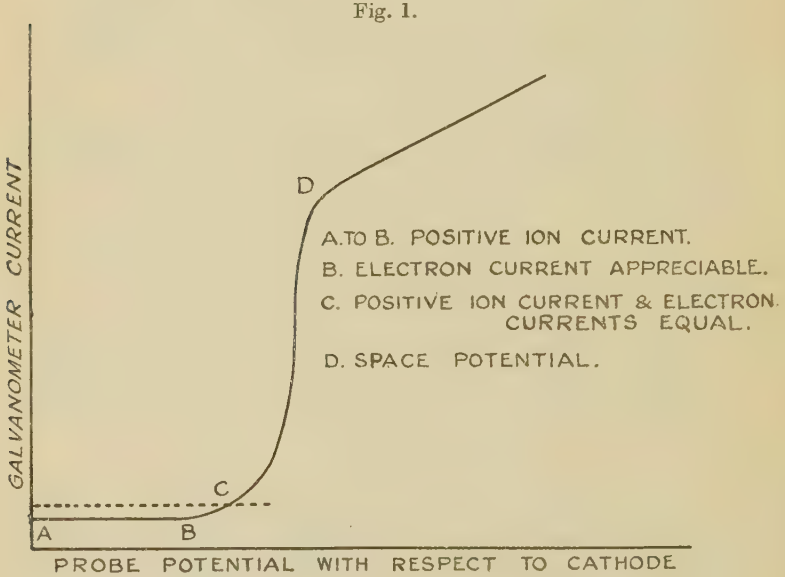
1. *Introduction.*

PROBES, introduced into an arc discharge for the measurement of space potential, arc gradient, cathode and anode falls, etc., have been used by many investigators, but the results obtained previous to 1924 were rendered invalid by the work of Langmuir and Mott-Smith §, who showed that under the conditions of no probe current the electron current exactly equals the positive ion current, so making the total current registered in the probe circuit zero. Under these conditions the probe potential differs considerably from the potential of the space in which the probe is situated. Thus, in order to measure the space potential along an arc discharge, it is not sufficient simply to immerse the

† 'Nature,' cxix. p. 280 (1927).
‡ Communicated by the Author.
§ Gen. Elect. Rev. xxvii. pp. 449, 538, 616, 762, 810 (1924).

probe in the arc and measure its potential, but the voltage-current characteristics of the probe must be studied, the probe being immersed in the arc and maintained at definite potentials with respect to one of the electrodes. If the probe potential is negative with respect to the space in which it is situated, a sheath of positive ions will form round the probe, and reflexion of those electrons will occur which, penetrating the sheath, have insufficient energy to pass to the probe against the field set up by the maintained potential on the probe. The sheath is of

Fig. 1.



Variation of probe current with probe potental.

such a thickness that the probe exerts no influence beyond its boundaries. On the other hand positive ions which penetrate the sheath will pass on to the probe, and the rate at which these ions strike it is a measure of the current in the circuit containing the probe. This positive ion current, as shown by AB, fig. 1, does not change much, as the probe potential is increased with respect to the cathode, but at B, since the probe potential is then slightly less than that of the space in which the probe is situated, fast-moving electrons penetrate the sheath and, striking the probe, give rise to an appreciable electron current.

The total current in the probe circuit is then the sum of the two—electron and positive ion currents.

At C the resultant probe current is zero, the two constituents, electron and positive ion currents, neutralizing. With probe potentials higher than that corresponding to C the electron current exceeds the positive ion current, and at this stage, if the electrons which come to the probe have a Maxwellian distribution of velocities, the following well-known relationships hold :

$$I=I_0\epsilon^{-eV/kT}, \quad I_0=Ne(kT/2\pi m)^{\frac{1}{2}}, \quad i=AI,$$

where I is the electron-current density to the probe whose potential is V with respect to the space potential, I_0 is the random electron-current density, N and T are the electron concentration and temperature respectively, A the probe area, e the electronic charge and m its mass, k the Boltzmann constant.

These considerations assume that the ions have a purely thermal distribution, whereas in practice there is always a drift of electrons and ions towards the electrodes, this drift current constituting the current through the external circuit. The ratio of random to drift currents in the low-pressure discharge is approximately unity, but Bramhall * has found that with the normal copper arc the actual ratio of drift to random currents may exceed 100. Thus the temperature motion of the electrons is relatively unimportant, and this result complicates the interpretation of the probe characteristics, so that the electron density and energy cannot be accurately determined from the probe measurements.

When the probe has a potential corresponding to C the sheath around the probe vanishes, and since the positive ions have a relatively low speed compared with that of the electrons, the positive ion current is so small as to be practically negligible, compared with the electron current.

From the above equations we have

$$\log i=\log AI_0-eV/kT=\log ANe(kT/2\pi m)^{\frac{1}{2}}-eV/kT,$$

which is the equation of the part CD of the curve. Thus if $\log i$ is plotted against V the part of the graph corresponding to CD gives a straight line, and the slope of this

* Phil. Mag. xiii. p. 682 (1932).

straight line portion is equal to e/kT. Since e/k is known, the temperature T can be calculated. The fact that this part of the curve is in practice straight proves that the electrons have a Maxwellian distribution of velocities, as assumed in the above equations.

The theory of the Langmuir probe method predicts that the current-voltage curve will show a change in slope when the probe potential equals that of the space, and it is on this basis that the bend of the curve at D is taken as an indication of the true potential of the space. Actually the electron current to the probe does not reach its maximum value at space potential, but only when the probe is slightly positive with respect to the space. This is attributed to the well-known phenomenon of electron reflexion from the probe, and the intervening curve depends on the distribution of velocities of the reflected electrons.

The probe method of measuring space potentials, as described above, is suitable for low-pressure discharges, but for arcs the bombardment of the probe by positive ions is so violent that it either melts or is raised to such a high temperature that it emits electrons, this electronic emission rendering the results invalid. Nottingham * modified the arrangements, swinging the probe by mechanical means at practically a constant velocity through the arc so rapidly that it did not heat up sufficiently to interfere with the experimental results, the current being measured by means of a ballistic galvanometer, suitably shunted. To calculate the space potential from these measurements it is not necessary to know the area of the arc stream swept over by the probe, and if v is the uniform velocity at which the probe is moved through the arc, r the radius of the arc through which the probe swings, i the current per unit length of the probe, and Q the quantity of electricity which flows through the galvanometer circuit, then

$$Q=\pi r^2 i/v.$$

Although it is recognized that this probe theory is not strictly applicable to the study of high-pressure arcs in which the electronic mean free path is so small that collisions in the ion sheath surrounding the probe can

* Journ. Franklin Inst. ccvi. p. 43 (1928).

no longer be neglected, yet the probe method still remains the most important method for elucidating the complicated processes occurring in normal arcs. It has been applied by many investigators for various metallic arcs, *e. g.*, Nottingham *, Compton and Lamar †, Killian ‡, Compton and Van Voorhis §, and Bramhall ||. It has also been used in the present work to determine the cathode fall of potential in the sodium arc.

The ordinary fine wire probe usually employed was found to be useless with this type of arc, as the wire quickly fused and, owing to mechanical difficulties, it was not possible to use the swinging probe method devised by Nottingham. After several experiments the probe finally employed consisted of a tungsten rod turned to a diameter of 1 mm. over a length of 10 mm. at one end. This end was then covered with a thin glass tube and the extremity ground flat. The remaining part of the tungsten rod was 6 mm. diameter, and by means of this device excessive heating of the probe was avoided. Measurements made with it were consistent, and showed no erratic values due to thermionic emission.

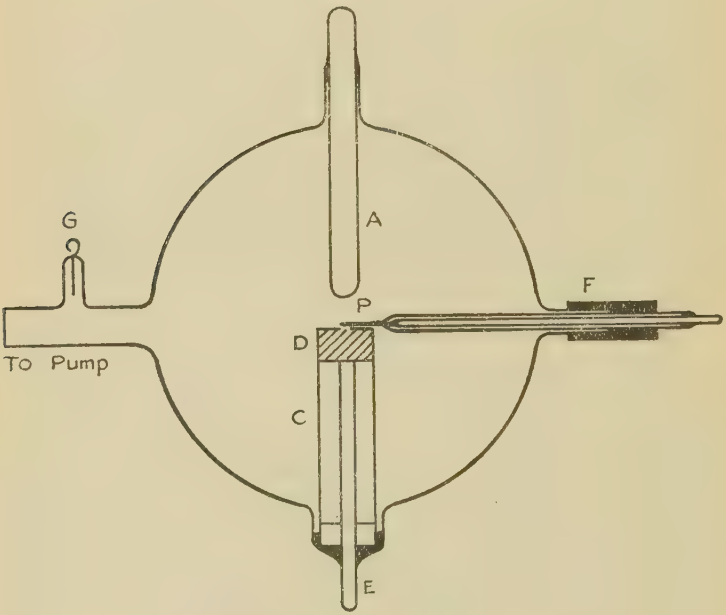
2. *Experimental.*

The type of arc tube utilized is shown diagramatically in fig. 2. The probe P could be moved over an exploring ranged of 10 mm. by means of the flexible rubber joint F. The sodium metal D, constituting the cathode, was packed at the end of a glass tube C, and maintained in position by means of an aluminimum plate to which was attached the aluminium electrode E. The end of the tube C was almost closed at the top, a small hole being left at which the sodium surface was exposed. By means of this device, originally used by Lamar and Compton ¶, wandering of the cathode bright spot was minimized. The anode A, made of iron tubing, was placed vertically above the cathode, G being an auxiliary electrode by means of which a transient electrical discharge could be passed through the tube to start the arc. With the large

* *Loc. cit.*
† Phys. Rev. xxxvii. p. 1069 (1931).
‡ *Ibid.* xxxi. p. 1122 (1928).
§ Nat. Acad. Sci. Proc. xiii. p. 336 (1927).
|| *Loc. cit.*
¶ *Loc. cit.*

glass bulb the vapour pressure and the temperature during the arc-discharge were both maintained within reasonable limits. Evaporation of the sodium, with consequent change in the position of its upper surface, was negligible during any short run. This was not unexpected, as in previous experiments it had been found that the rate of evaporation of a sodium cathode is small—a fact that suggests the temperature of such a cathode, when the arc is passing, is much lower than that generally

Fig. 2.



Sodium arc tube.

assumed. No trouble was experienced through condensation of the sodium vapour on the exposed part of the probe while the arc was passing, but condensation did occur on the exposed surface of P after the arc ceased, and it was necessary to remove P and clean the surface after each discharge. The positions of D and P could be adjusted, however, after such operations so as to insure that the actual distances between A, P, and D remained the same during any particular set of experiments.

The experimental procedure was to exhaust the discharge tube, apply a potential difference across E and A, and pass a momentary electrical discharge from a transformer between G and E. This started the arc, and with a known potential difference between P and E the reading of the galvanometer in the probe circuit was taken. Without interrupting the arc various potential differences were applied between the probe and the cathode and the corresponding galvanometer currents observed. The observations were made as rapidly as possible to avoid any change in the distance between the probe and the upper surface of the sodium metal—this distance being observed before and after the passing of each arc-discharge. In all cases the potential of P was, of course, positive with respect to the cathode.

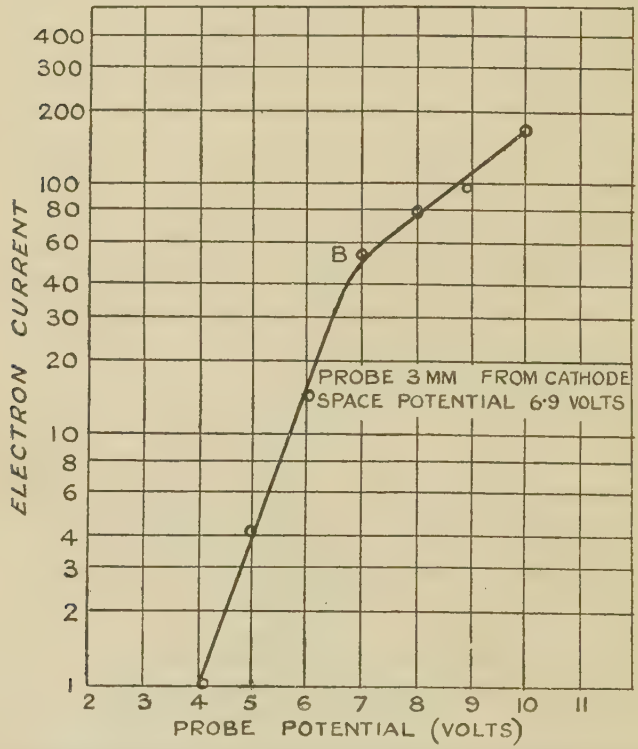
At first difficulties were encountered from the spurting of the sodium metal from D as it melted, owing to the escape of the trapped air within the mass of the metal. These difficulties were removed by drilling holes in the aluminium plate supporting the sodium and in the glass tube C. The initial film of oxide on the surface of the sodium presented no trouble as, after the arc had started, this film disappeared and the surface became quite clean.

The distance between the lower end of A and the sodium surface was 14 mm., and the arc-current was obtained by means of 110 volts D.C., with an external resistance in the circuit to adjust the current to any desired value. Fig. 3 shows the results obtained when the logarithm of the galvanometer current in the probe circuit was plotted against the potential of the probe, the centre of the probe being 3 mm. from the cathode. The overall voltage, *i. e.*, the potential difference between the cathode and anode, was 11 volts, and the arc-current 5·0 amperes. The potential of the space 3 mm. from the cathode thus corresponds to the point B, just above the termination of the straight portion of the curve, viz., 6·9 volts.

Proceeding in this manner, the space potential was determined at various points along the arc-stream, and the results are shown in fig. 4, which indicates the distribution of space potential along the arc from the cathode, taken as zero, to the anode at 11·0 volts. Straight-line extrapolations through the observed points gives a cathode fall of about 5·0 volts and an anode fall of 0·6 volt. The value of the cathode fall agrees approximately with

that of the experimentally determined ionization potential of sodium vapour, viz., 5·1 volts, and extends the general results found by other experimenters that the cathode fall is equal to, or slightly greater than, the ionization potential of the vapour actively involved. This observation was corroborated spectroscopically by the fact that the entire arc-spectrum could be observed at the surface

Fig. 3.



Curve showing variation of electron current with probe potential.

of the cathode when the image of the arc was carefully focussed on the slit of a spectrograph.

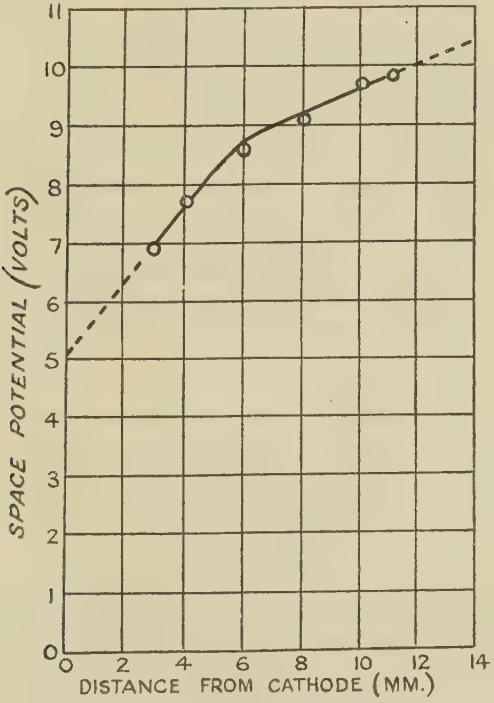
It has been shown by the author * that the residual gases in a discharge tube, as well as the active vapour from the electrodes, are involved when an arc passes between metallic poles, but after the sodium arc has been

* Phil. Mag. xiv. p. 788 (1932).

burning for some time spectroscopic observations indicate that the sodium vapour is the all important active component in the arc-discharge, and, therefore, the cathode fall corresponding to the ionization potential of sodium vapour is not unexpected.

The cathode fall is not, however, necessarily the ionizing potential of the metallic vapour from the cathode;

Fig. 4.



Variation of potential along the sodium arc.

for example, the cathode fall for a tungsten arc is 16·2 volts *, whereas Anderson and Kretchmar † found that the total potential difference across a tungsten arc was 14 volts at 12 amperes. In addition, the introduction of the probe causes a small increase in the overall voltage. This is probably due to a local cooling of the gases near

* Phys. Rev. xxi. p. 266 (1923).
† *Ibid.* xxvi. p. 33 (1925).

the probe. The rate of recombination of the ions is accordingly increased, and this necessitates an increased rate of ion production which can only be brought about by an increase in the potential gradient near the probe.

Unfortunately, although it is a comparatively simple matter to measure the value of the cathode fall, the determination of the distance across which this fall occurs is very difficult. It has never been actually measured, but Lamar and Compton * have calculated it to be less than $1{\cdot}76\times10^{-4}$ cm. in the mercury arc. They assume, however, that the random electron current is larger than the drift current, an assumption that does not appear to be entirely justified.

The determination of the cathode-fall space is a matter of importance since, if it can be found, the electric field at the cathode surface can then be calculated. This electric field may be of sufficient strength to draw out electrons from the cathode surface, a theory of the arc proposed by Langmuir. The cathode surface, however, is probably comparatively uneven, and it seems impossible to determine the extent of the cathode-fall space.

LI. *The Effect of small Water Additions upon the Conductivity of Electrolytes in non-aqueous Solvents.* *By* O. L. Hughes *and* Sir Harold Hartley, *F.R.S.*†

Introduction.

THE Debye-Hückel-Onsager[1] theory explaining the behaviour of strong electrolytes in solution is found to be in good agreement with the experimental results in water. The observed slopes of the electrical conductivity curves agree well with the theoretical values[2]. On passing from water to other solvents of lower dielectric constant deviations from perfect behaviour appear, which are due to the association of the ions to form molecules, and these deviations are greater the smaller the dielectric constant, that is to say, the larger the interionic forces. The problem is not, however, so simple as to be entirely explained by the fall in dielectric constant, and measurements in a variety of solvents have indicated that ion solvation is of great

* *Loc. cit.*

† Communicated by the Authors.

importance in modifying the behaviour of an electrolyte in solution. For example, in methyl-alcohol all uni-univalent salts are strong electrolytes, but in nitromethane, a solvent of greater dielectric constant, a large number of these salts are weak electrolytes. This difference in behaviour is explained by the fact that nitromethane is electronically only a donor solvent, and is not capable of solvating the anions chemically. The pseudo-acid form of nitromethane should be capable of solvating anions, but it is not present to any great extent.

There are two outstanding problems in the field of electrical conductivity: one is the relation between the ion and the solvent molecule, the problem of solvation; the other is the relation between the ions themselves, the problem of ion association. It appears that these two factors, on which the behaviour of an electrolyte in solution ultimately depends, operate in opposite directions, which makes their study difficult. Solvation is of two kinds—there is the so-called chemical solvation when there is a coordinate link between the ion and the solvent molecule (Sidgwick[3]); and there is the type of solvation discussed by Born and Schmick[4], in which the solvent molecules are orientated around the ion by electrical dipole forces. The work of Washburn[5], Remy[6], and Barborovsky[7] has given an indication of the number of water molecules associated with various ions. Ulich[8] has estimated the degree of solvation of certain ions from the difference between the size of the ion in the lattice and its size in solution, as calculated from the mobility by means of Stokes's Law. The information derived from conductivity data is rather more uncertain, because both factors are operating at once, and a clear differentiation between them is impossible. One method of attack is to examine the effect of small water additions upon electrical conductivities in non-aqueous solvents.

Ulich[9] has pointed out that a small amount of water, say 0·5 per cent., does not modify the physical properties of a non-aqueous solvent to any great extent; in fact, the viscosity is the only property which is altered considerably. The alteration in the conductivity varies greatly with the nature of the electrolyte. The conductivity of a strong salt is in general reduced, and the depression is of the same order as the increase in the viscosity. The conductivity of a weak salt is greatly increased, presumably owing to an increase in the degree of dissociation, and this increase is of a quite different order from the decrease observed in the case of strong salts. Traces of water affect strong acids in a manner

which is peculiar to the solvent; in the alcohols, in which the hydrogen ion has an abnormally high mobility, a large decrease of conductivity is observed, which is of a higher order than the viscosity change involved; in non-hydroxylic solvents, in which the hydrogen ion has a normal mobility, a strong acid behaves just like a strong salt, and the change observed is of the same order as the viscosity change. Weak acids exhibit much the same behaviour as weak salts in both types of solvent.

Goldschmidt[10] and his co-workers have found that, whereas in the case of sodium chloride in methyl alcohol the presence of 1 per cent. of water produced a change of 2·4 per cent. in the conductivity at 0·0007 N., the same amount of water reduced the conductivity of hydrobromic acid by 25 per cent. at the same concentration. The change in viscosity was found to be 4 per cent. They also found that the changes produced in the conductivities at infinite dilution were larger than those produced at 0·0007 N., owing to the slopes of the conductivity curves having been reduced. For weak acids in methyl and ethyl alcohol they found large increases in the conductivity, 1 per cent. of water producing an increase of about 50 per cent. at concentrations in the neighbourhood of 0·001 N. Similar results have been obtained by Murray-Rust and Hartley[11] for acids, and Unmack, Murray-Rust, and Hartley[12] for salts. In non-hydroxylic solvents, in which the hydrion has a normal mobility, strong acids behave in the same way as strong salts; this has been demonstrated for perchloric acid in acetone by Ross Kane and Hartley[13], and for the same acid in nitromethane by Wright, Murray-Rust, and Hartley[14]. They have also shown that water has a very large effect on the conductivity of weak electrolytes in these solvents.

With the exception of the work of Goldschmidt, the water additions so far made have been confined to single-point observations, a small quantity of water having been added to the solution at the end of a series of measurements in the dry solvent. This is liable to be a dangerous proceeding, as the effect of the water addition may be masked, or at any rate obscured, by any heat change on mixing the solvents, for solutions take some time to come to temperature equilibrium. The present work was planned to investigate the effect of varying amounts of water on the conductivity over a range of concentration, with a view to studying:

(1) The relation between the amount of water added and the effect produced.

(2) The effect of water on the slope of the conductivity curves for strong electrolytes.

(3) The effect of water on the dissociation of weak salts. It was hoped to obtain from such measurements information as to the effect produced by these water additions firstly on the solvation of the ions, and secondly on the degree of association.

Preparation of Materials.

Solvents.

The preparation of solvents was carried out according to the methods of the following investigators: methyl alcohol, Hartley and Raikes[15]; ethyl alcohol, Copley, Murray-Rust, and Hartley[16]; acetone, Ross Kane and Hartley[13]; water, Bourdillon[17].

Salts.

In many cases specimens were available in the laboratory, and certain other salts were prepared according to the directions of previous workers.

Picric acid was prepared by Mr. T. H. Mead by crystallizing the B.D.H. product twice from benzene and once from ethyl alcohol. It was dried in a vacuum desiccator over phosphorus pentoxide.

Silver picrate was prepared by Mr. Ross Kane from picric acid and silver oxide; it was recrystallized from acetone, and dried at 80° C. in a vacuum oven over phosphorus pentoxide.

Density and Viscosity Measurements.

The density values of Goldschmidt and Aarflot were employed for water-methyl alcohol and water-ethyl alcohol

$$\eta^{25^\circ\,\mathrm{C.}}_{\mathrm{Water.}} = 0{\cdot}00895.$$

	Percentage of water.	Viscosity.
Methyl alcohol	No water.	0·00545
	0·245	0·00549
	0·478	$0{\cdot}00554_5$
	1·080	0·00567
Ethyl alcohol	No water.	0·01078
	0·494	0·01099
	1·031	0·01124
Acetone	No water.	$0{\cdot}00307_5$
	0·250	$0{\cdot}00309_5$
	0·493	0·00312
	0·769	0·00314
	1·020	0·00316

mixtures. Viscosity measurements were made for the various mixtures used.

There was considerable doubt as to the density of acetone, and no measurements had been made of the density of acetone containing small quantities of water, so a series of measurements was made with an Ostwald pycnometer.

Percentage of water.	Δ^{25} referred to water at 4° C.
No water	0·78501 0·78505
0·250	0·78590 0·78584
0·493	0·78656 0·78646
0·996	0·78797 0·78807

The density values for acetone-methyl alcohol mixtures and for acetone-ethyl alcohol mixtures were taken from the results of Jones and Bingham [18].

Experimental Procedure.

All solutions and mixtures of solvents were made up by weight. A solution of perchloric acid in acetone was made in the manner described by Ross Kane [13] by mixing a solution of silver perchlorate and a solution of hydrogen chloride. The precipitate of silver chloride was allowed to settle and the solution drawn off.

The conductivity measurements were made by the method of Frazer and Hartley [19], as modified by Murray-Rust and Hartley [11]. The conductivity of the pure solvent was first measured, and additions of solution were made from a weight pipette. All measurements were made at 25·00 ± 0·01° C. The conductivity of the pure solvent was subtracted from the conductivity of the solution, except in the case of perchloric acid when no correction was made, because it was assumed that the conductivity of traces of impurities such as carbon dioxde would be suppressed in the presence of a strong acid (Wynne-Jones) [20]. The constants of the cells employed were based on the results obtained by Frazer and Hartley for potassium chloride in methyl alcohol at 25° C. The constant of the cell used in the measurements in the alcohols was 0·03826, while that of the acetone cell was $0{\cdot}1513_4$.

Results.

The results are given in the following tables. "k" is the specific conductivity of the solvent employed, given in reciprocal megohms. In the cases where water additions are made, the amount of water added is given. The variation of electrical conductivity with concentration is given for strong electrolytes by the equation :—

$$\Lambda_c = \Lambda_0 - x\sqrt{c}$$

and values for Λ_0, the conductivity at infinite dilution, and "x," the slope of the conductivity curve, are given. The concentrations of the various solutions, expressed in gram equivalents per litre, and multiplied by 10^4, are given in the first column. The observed conductivities are given in the second column, and the differences between the observed values and the values calculated from the equation appear in the last column.

Potassium Chloride in Methyl Alcohol.

In dry methyl alcohol, $\Lambda_c = 105{\cdot}0_7 - 261\sqrt{c}$, and this equation was employed for determining the cell constant.

	$c\times10^4$.	Λ_c.	Diff.
k=0·260.	1·0565	102·02	−0·01
0·245 p. c. water.	2·0916	101·00	+0·03
Λ_0=104·65.	4·4658	99·31	+0·02
x=255.	7·6005	97·63	+0·01
	10·8664	96·25	0·00
	14·6356	94·80	−0·09
k=0·114.	1·2165	101·47	+0·11
0·490 p. c. water.	2·4907	100·13	+0·01
Λ_0=104·21.	4·5930	98·66	0·00
x=259.	7·7887	96 94	−0·04
	10·8586	95·72	+0·05
	14·962	94·14	−0·05
k=0·112.	1·2155	100·23	+0·02
1·026 p. c. water.	2·7492	98·78	0·00
Λ_0=103·04.	5·7384	96·85	−0·04
x=257.	10·2737	94·82	+0·02
	15·3488	92·95	−0·02

Silver Nitrate in Methyl Alcohol.

	$c\times10^4$.	Λ_c.	Diff.
k=0·063.	1·5891	106·94	−0·36
No water.	2·6274	105·54	−0·15
Λ_0=112·95.	4·5327	103·43	+0·02
x=448.	7·7660	100·56	+0·20
	10·3086	98·61	+0·04
	13·9503	96·20	0·00

Silver Nitrate in Methyl Alcohol (*cont.*).

	$c\times10^4$.	Λ_c.	Diff.
k=0·068. 0·536 p. c. water. Λ_0=111·56. x=430.	1·6061	105·78	−0·33
	2·6729	104·39	−0·18
	4·3457	102·58	−0·01
	7·1800	100·04	0·00
	9·6718	98·20	+0·01
	12·8300	96·16	0·00
k=0·092. 1·081 p. c. water. Λ_0=109·93. x=410.	1·4746	104·50	−0·45
	2·4169	103·36	−0·20
	4·2835	101·41	−0·03
	7·0857	99·02	+0 01
	9·6415	97·20	0·00
	13·2732	94·93	+0·01

Tetraethylammonium Perchlorate in Methyl Alcohol.

	$c\times10^4$.	Λ_c.	Diff.
k=0·098. No water. Λ_0=133·20. x=438.	0·66688	129·21	−0·41
	1·3114	128·06	−0·12
	2·4288	126·41	+0·03
	4·3545	124·17	+0·11
	5·8605	122·66	+0·06
	8·3087	120·46	−0·12
k=0·119. 0·267 p. c. water. Λ_0=132·10. x=425.	0·48114	128·89	−0·26
	0·93712	127·90	−0·08
	1·9873	126·14	+0·01
	2·8594	124·93	+0·02
	4·0564	123·58	+0·04
	5·8736	121·78	−0·02
k=0·140. 0·453 p. c. water. Λ_0=131·30. x=421.	0·9275	126·99	−0·26
	1·7267	125·73	−0·04
	3·0819	123·93	+0·02
	5·2999	121·65	+0·04
	7·2856	119·91	−0·02
	10·0858	117·73	−0·20

Barium Perchlorate in Methyl Alcohol.

	$c\times10^4$.	Λ_c.	Diff.
k=0·086. No water. Λ_0=131·78. x=895.	1·1394	122·44	+0·21
	2·2927	118·23	0·00
	4·3128	112·94	−0·26
	7·8020	106·61	−0·17
	10·5302	102·84	+0·12
	15·1672	98·08	+1·06

Barium Perchlorate in Methyl Alcohol (*cont.*).

	$c\times10^4$.	Λ_c.	Diff.
k=0·196.	1·5541	119·58	+0·04
0·459 p. c. water.	2·7995	115·69	−0·07
Λ_0=130·60.	5·1044	110·56	0·00
x=887	8·5485	105·09	+0·07
	12·5887	100 33	+1·20
	17·5049	96·00	+2·59

Tetraethylammonium Perchlorate in Ethyl Alcohol.

	$c\times10^4$.	Λ_c.	Diff.
k=0·022.	0·44272	59·50	−0·35
No water.	0·84625	58·66	−0·03
Λ_0=62·65.	1·7607	56·99	+0·05
x=430.	2·5231	55·84	+0·02
	3·4511	54·62	−0·04
	4·9481	52·95	−0·14
k=0·035.	0·44265	58·87	−0·36
0·494 p. c. water.	0·82143	58·10	−0·11
Λ_0=62·02	1·7214	56·54	+0 03
x=420.	2·3929	55·55	+0·03
	3·3089	54·36	−0·02
	4·7377	52·57	−0·31
k=0·028.	0·57530	57·93	−0·30
1·031 p. c. water.	1·0951	56·99	−0·06
Λ_0=61·34.	2·1254	55·39	+0·01
x=410.	3·0042	54·27	+0·05
	4·2126	52·93	+0·01
	6·2176	51·08	−0·14

Tetraethylammonium Perchlorate in Acetone.

	$c\times10^4$.	Λ_c.	Diff.
k=0·128.	1·2461	196·1	+0·2
No water.	2·3068	191·4	+0·1
Λ_0=208·5.	4·1847	185·2	−0·2
x=1130.	7·7707	176·5	−0·4
	10·5930	171·6	−0·1
	14·2184	166·0	+0·1
k=0·076.	1·3635	194·1	+0·2
0·250 p. c. water.	3·0142	187·7	+0·1
Λ_0=206·9.	6·3798	178·7	−0·1
x=1113.	11·9945	168·4	0·0
	16·3474	162·7	−0·8
	21·872	157·1	+2·3

Tetraethylammonium Perchlorate in Acetone (*cont.*).

	$c\times10^4$.	Λ_c.	Diff.
k=0·096. 0·493 p. c. water. Λ_0=205·7. x=1085.	1·1178	194·2	0·0
	2·2783	189·3	0·0
	4·6043	182·3	−0·1
	8·9158	173·2	−0·1
	11·7558	168·7	+0·2
	16·2429	161·4	−0·6
k=0·210. 0·996 p. c. water. Λ_0=203·1. x=1053.	1·5742	189·8	−0·1
	2·9646	184·9	−0·1
	6·3834	176·3	−0·2
	11·1921	168·1	+0·2
	15·1770	162 5	+0·4
	21·621	156·0	+1·9

Potassium Iodide in Acetone.

	$c\times10^4$.	Λ_c.	Diff.
k=0·018. No water. Λ_0=192·1. x=980.	1·3235	180·6	−0·2
	2·3906	177·1	+0·1
	4·5701	171·3	+0·2
	8·5866	163·3	−0·1
	11·3236	159·0	−0·1
	15·900	152·9	−0·1
k=0·051. 0·769 p. c. water. Λ_0=187·1. k=910.	1·6182	175·3	−0·2
	3·0781	171·3	+0·2
	5·8882	164·9	−0·1
	12·6284	154·4	0·0
	16·7585	149·6	−0·2
	23·267	143·4	+0·2

Silver Nitrate in Acetone.

	$c\times10^4$.	Λ_c.
k=0·071. No water.	0·8216	37·85
	1·7781	27·65
	3·1893	22·03
	4·6267	19·28
	6·2325	17·48
	8·6547	15·78
k=0·118. 0·267 p. c. water.	1·0363	38·80
	1·8359	30·84
	3·8672	23·06
	5·2572	20·61
	7·5296	18·21

Silver Nitrate in Acetone (*cont.*).

	$c\times10^4$.	Λ_c.
k=0·140. 1·015 p. c. water.	0·9203	55·37
	1·8559	42·44
	3·5730	32·75
	5·2680	28·09
	7·1030	25·10
	10·1068	22·07
k=0·118. 0·961 p. c. MeOH.	0·9600	42·92
	1·9804	31·93
	4·0381	23·97
	5·4969	21·35
	8·2796	18·47
	11·9493	16·40
	11·90 *	20·8
	* 0·67 p. c. water added.	
k=0·098. 1·957 p. c. MeOH.	1·0727	45·61
	1·9969	35·18
	3·6248	27·50
	5·2270	23·78
	7·0364	21·23
	9·8417	18·85
	9·74 *	20·21
	* 1·00 p. c. EtOH added.	
k=0·425. 1·916 p. c. EtOH.	1·0582	41·42
	2·0472	30·92
	4·1331	23·19
	5·6456	20·60
	7·8317	18·31
	11·2034	16·29
	11·16 *	20·8
	* 0·67 p. c. water added.	
k=0·254. 3·984 p. c. EtOH.	1·1070	47·27
	2·0793	36·66
	4·5537	26·65
	6·4709	22·97
	9·1569	20·34
	11·6427	18·64
	11·52 *	21·67
	* 1·07 p. c. MeOH added.	

Perchloric Acid in Acetone.

	$c\times10^4$.	Λ_c.	Diff.
k=0·043. No water. Λ_0=207·1. k=1400.	1·1955	192·5	+0·7
	2·2075	186·5	+0·2
	4·1797	178·1	−0·4
	7·1756	169·0	−0·6
	10·5380	161·3	−0·3
	14·3470	154·6	+0·6

Perchloric Acid in Acetone (*cont.*).

	$c\times10^4$.	Λc.	Diff.
$k=0{\cdot}114$.	0·66028	194·6	+3·2
0·530 p. c. water.	1·4992	187·2	+1·1
$\Lambda_0=202{\cdot}0$.	2·8249	180·4	+0·2
$x=1295$.	5·2578	171·6	−0·6
	7·3694	166·1	−0·6
	9·8534	161·1	+1·1
$k=0{\cdot}062$.	0·76501	186·0	+0·7
0·939 p. c. water.	1·7694	180·0	−0·1
$\Lambda_0=195{\cdot}2$.	3·2474	174·4	−0·3
$x=1135$.	5 6759	167·6	−0·6
	8·5560	161·6	−0·4
	12·460	155·5	+0·4

Picric Acid in Acetone.

	$c\times10^4$.	Λc.
$\kappa=0{\cdot}026$.	1·0373	6·65
No water.	2·1053	3·70
	4·0202	2·27
	6·8914	1·56
	9·7951	1·23
	12·7714	1·03
$k=0{\cdot}058$.	0·9570	15·82
0·504 p. c. water.	1·9705	9·69
	3·6794	6·51
	7·0539	4·40
	10·0204	3·58
	13·1936	3·07
$k=0{\cdot}061$.	1·0890	23·69
1·007 p. c. water.	2·0852	17·15
	3·9126	12·54
	5·9467	10·97
	8·9328	8·76
	12·5504	7·30

Silver Picrate in Acetone.

	$c\times10^4$.	Λc.	Diff.
$k=0{\cdot}048$.	0·5516	155·9	0·0
No water.	0·9712	147·3	0·0
$\Lambda_0=181{\cdot}8$.	1·9471	133·1	+0·1
$x=3500$.	3·8764	115·7	+2·5
	5·6104	105·6	+6·7
	7·3090	98·4	+11·2

Silver Picrate in Acetone (*cont.*).

	$c\times10^4$.	Λ_c.	Diff.
$k=0{\cdot}151$.	·7389	150·0	−0·2
0·481 p. c. water.	1·1848	143·9	+0·1
$\Lambda_0=174{\cdot}3$.	2·0448	134·2	−0·1
$x=2800$	3·3940	123·0	+0·3
	4·6758	115·0	+1·2
	6·1666	108·2	+3·4
$k=0{\cdot}075$.	0·4567	155·5	0·0
0·996 p. c. water.	1·0188	148·7	+0·5
$\Lambda_0=170{\cdot}4$.	2·0002	139·0	−0·3
$x=2200$.	3·5602	127·9	−1·0
	5·1298	119·8	−0·8
	6·9991	112·4	+0·2

Potassium Thiocyanate in Acetone.

	$c\times10^4$.	Λ_c.	Diff.
$k=0{\cdot}057$.	1·2244	191·5	−1·2
No water.	2·3873	184·2	+0·1
$\Lambda_0=207{\cdot}3$.	4·8150	174·3	−0·1
$x=1500$.	8·1326	164·4	−0·1
	11·2851	157·2	+0·2
	15·4280	149·4	+1·0
$k=0{\cdot}255$.	1·7502	183·8	0·0
0·516 p. c. water.	3·2563	177·3	0·0
$\Lambda_0=201{\cdot}7$.	5·6228	169·5	−0·2
$x=1350$.	8·5732	162·0	−0·2
	11·9137	155·3	+0·2
	15·9714	148·9	+1·2
$k=0{\cdot}270$.	2·0774	178·8	+0·1
1·011 p. c. water.	3·5806	173·2	0·0
$\Lambda_0=196{\cdot}4$.	5·8612	166·6	−0·1
$x=1225$.	9·0300	159·4	−0·2
	12·1031	153·9	+0·1
	15·8613	148·3	+0·7

Sodium Thiocyanate in Acetone.

	$c\times10^4$.	Λ_c.	Diff.
$k=0{\cdot}096$.	1·4170	178·0	+0·2
No water.	2·6780	166·6	+0·6
$\Lambda_0=202{\cdot}0$.	5·0185	152·0	−0·7
$x=2200$.	8·3907	138·1	−0·2
	11·6375	128·6	+1·6
	15·8476	119·5	+5·1

Sodium Thiocyanate in Acetone (*cont.*).

	$c\times10^4$.	Λ_c.	Diff.
$k=0{\cdot}090$.	1·4417	175·2	+0·2
0·659 p. c. water.	2·9960	164·9	−0·3
$\Delta_0=197{\cdot}0$.	5·7903	152·3	−0·5
$x=1835$.	9·1697	141·6	+0·2
	12·3053	134·2	+1·6
	16·9349	125·9	+3·4
$k=0{\cdot}240$.	1·6086	172·2	+0·1
0·993 p. c. water.	3·0949	163·5	−0·2
$\Lambda_0=194{\cdot}0$.	6·0698	151·1	−0·4
$x=1725$.	8·7672	143·0	+0·1
	11·8247	135·8	+1·1
	16·4510	127·5	+3·5

Ammonium Thiocyanate in Acetone.

	$c\times10^4$.	Λ_c.	Diff.
$k=0{\cdot}066$.	1·0636	171·5	+0·3
No water.	2·0249	155·5	+0·6
$\Lambda_0=214$.	3·9922	135·3	+4·2
$x=4150$.	7·3321	115·6	+14·0
	10·8310	102·9	—
	15·2877	92·1	—
$k=0{\cdot}083$.	1·3043	172·1	+0·5
0·483 p. c. water.	2·7470	153·4	−0·4
$\Lambda_0=211$.	4·8501	136·9	+1·9
$x=3450$.	8·1244	120·7	+8·0
	13·9335	104·2	—
$k=0{\cdot}029$.	1·2176	174·7	−0·5
0·970 p. c. water.	2·5008	160·3	−0·7
$\Lambda_0=208$.	4·5002	145·3	+0·3
$x=2970$.	7·6166	129·9	+3·9
	10·8912	119·0	—
	15·1558	108·9	—

Lithium Thiocyanate in Acetone.

	$c\times10^4$.	Λ_c.
$k=0{\cdot}021$	1·1922	118·2
No water.	2·4156	95·71
	4·2332	78·91
	6·9759	65·69
	10·1130	56·88
	13·6596	50·48

Lithium Thiocyanate in Acetone (*cont.*).

	$c\times10^4$.	Λc.
k=0·026 No water.	1·1997	117·7
	2·5555	93·79
	4·5790	76·70
	7·6421	63·40
	11·0182	54·92
	14·4713	49·23
k=0·060 0·521 p. c. water	1·2352	141·1
	2·3657	122·3
	4·3803	103·6
	7·5328	87·54
	10·5156	78·19
	14·1728	70·32
k=0·033 1·034 p. c. water.	1·2573	149·1
	2·1751	136·3
	4·1656	119·0
	7·2069	103·2
	10·2144	93·05
	13·4211	85·55

Experiments with Larger Water Additions.

Potassium Chloride in Methyl Alcohol.
[Conc. 0·001142 N.]

	Percentage of water.	Λc.	Diff.
k=0·057.	No water.	96·30	0·00
	0·824	94·94	0·00
	1·642	93·62	+0·03
	2·557	92·29	+0·21
	3·702	90·75	+0·56
	4·487	89·70	+0·80
	6·228	87·72	+1·70
k=0·040.	No water.	96·29	0·00
	1·564	93·86	+0·15
	3·458	91·09	+0·51
	6·419	87·48	+0·79
	9·049	84·85	+3·49
	13·459	82·02	+6·29

The last column gives the differences between the observed values of Λ_c and those calculated from the equation

$$\Lambda_c^{\text{No water.}} = \Lambda_c^{\text{Obs.}} + 1{\cdot}650\ \text{W},$$

where W is the percentage of water added.

Comments on Individual Determinations.

Tetraethylammonium perchlorate.

This salt does not give straight line relations between Λ_c and $\sqrt{c}$ either in methyl or in ethyl alcohol, the departure being particularly noticeable in the dilute range. It is significant that in both alcohols this salt has an abnormally steep slope, while in acetone, in which it has a more normal slope, there is no falling away in the dilute range.

Silver Nitrate.

The determinations in methyl alcohol exhibit the same curvature in the dilute range, and this is another case of a salt with a large slope. It appears to be quite a general phenomenon for electrolytes with slopes which are much greater than the Onsager value. It follows that mobility values for such salts obtained by straight-line extrapolation tend to be too high.

The effect of small additions of methyl and ethyl alcohol to acetone solutions of this salt was measured, and was found to be smaller than that produced by water. At the end of some of the determinations (4 out of 7) further small additions were made in order to test the magnitude of the effect produced and to see whether it is approximately identical with that produced by the same addition if made to a pure solvent.

Perchloric Acid in Acetone.

Owing to the difficulties of making up the solutions and to the possible errors arising from the process of mixing, these values are subject to rather higher errors than those for ordinary determinations. Solutions of perchloric acid go brown on standing, as Ross Kane [13] has found, and the hydrion concentration decreases, but since the solutions were used as soon as they were made up, and since there was no colour change during the time of the experiments, the error involved is probably small. It was noticeable that solutions containing small amounts of water seemed more stable than solutions in pure acetone. The results indicate that perchloric acid behaves exactly like a strong salt in acetone, which is to be expected on account of the normal mobility of the hydrion in this solvent.

Ammonium Thiocyanate in Acetone.

The values found in dry acetone lie on a smooth curve, but water additions bring about a marked straightening of

the curve. The differences, which give the departure of Λ_c from the value required by the square-root relation, are considerably smaller in solutions containing water.

Potassium Chloride in Methyl Alcohol containing larger amounts of water.

In the two experiments performed a different procedure was adopted from that normally employed. Additions of water were made to solutions of potassium chloride, and the change of conductivity measured. The calculation of the equivalent conductivity is complicated by the fact that the solvent correction is different for each measurement, and is not known. An approximate value was computed by assuming that there was a uniform change in the conductivity of the solvent over the whole range of pure alcohol to pure water. The error involved in such an assumption is not likely to be large.

In certain cases the results in the dry solvents can be compared with the work of previous investigators. In every case the agreement is satisfactory. The determinations in question are in methyl alcohol, silver nitrate[19], tetraethylammonium perchlorate[21], and barium perchlorate[22]; in ethyl alcohol, tetraethylammonium perchlorate[23]; and in acetone, tetraethylammonium perchlorate, potassium iodide, silver nitrate, perchloric acid, potassium thiocyanate, sodium thiocyanate, and ammonium thiocyanate[13].

Discussion of Results.

Before discussing measurements made in solution it is interesting to examine the results obtained by Tyndall[24]

Mobilities of Alkali Metal Ions in. cm./sec. per volt/cm.

	In argon.	In methyl alcohol.
Sodium.....	3·21	$4{\cdot}73\times10^{-4}$
Potassium	2·77	$5{\cdot}57\times10^{-4}$
Rubidium	2·37	$5{\cdot}95\times10^{-4}$,
Cæsium.....	2·23	$6{\cdot}48\times10^{-4}$,

for the mobilities of ions in gases. The mobilities of ions in a rare gas medium provide the simplest case, as there is no association and solvation is reduced to a minimum. Even in the rare gases the ions tend to have molecules held to

them by a dipole effect, but this is considerably smaller than in more polarizable gases such as nitrogen. When nearly free from disturbing effects the ion mobilities are in the order expected, the lightest ions moving fastest and the heaviest ions slowest.

When an electrolyte is dissolved in a solvent conditions are much more complex, for the behaviour is complicated by ion association and ion solvation. The solvation of the ions is so considerable that the order of the mobilities in

TABLE I.

Percentage Change in Λ_c with varying Proportions of Water.

	Methyl alcohol (conc. 0·0015 N).			Ethyl alcohol (0·001 N).
Water.	KCl.	$AgNO_3$.	NEt_4ClO_4.	NEt_4ClO_4.
per cent.				
¼	−0·4		−0·5	
½	−0·9	−0·6	−1·1	−0·6
1	−2·1	−1·4		−1·3

Acetone (0·001 N).

Water.	NEt_4ClO_4.	$HClO_4$.	KCNS.	AgPic.	NaCNS.
per cent.					
¼	−0·5				
½	−0·8	−1·2	−1·0	+2·2	+3·2
1	−1·5	−2·4	−2·0	+4·6	+6·0

Water.	$AgNO_3$.	NH_4CNS.	LiCNS.	Picric acid.
per cent.				
¼	+11·5			
½		+ 8	+35	+250
1	+48	+15	+65	+550

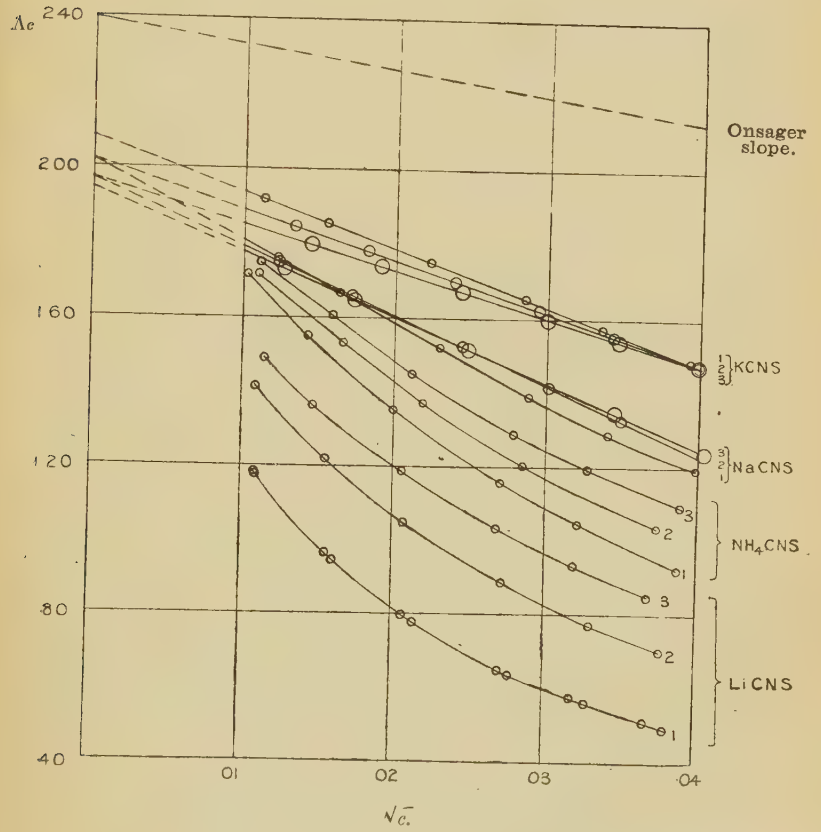
methyl alcohol of the alkali metal series is reversed owing to the higher degree of solvation of the lower members.

The most important conclusion which emerges from our measurements is the proportionality which exists between the amount of water added and the effect produced on the conductivity (see Table I.). Taking advantage of this proportionality, the changes in the conductivity have been adjusted to correspond to the addition of ¼ per cent., ½ per cent., and 1 per cent. of water.

It should be noted that the addition of 1 per cent. of water produces a 4 per cent. increase in viscosity in both the alcohols, while in acetone the same amount of water produces an increase of 2·7 per cent. in the viscosity.

The conductivity results for the thiocyanates in acetone are represented in fig. 1 in order to show how the effect of

Fig. 1.



Thiocyanates in acetone.

adding water to the solvent depends on the weakness of the electrolyte.

The curves numbered 1 for each electrolyte refer to pure solvent, while those numbered 2 and 3 refer to experiments with solvent containing increasing amounts of water, the data for which are given in the tables.

It is necessary to bear in mind the proportion of solvent molecules to ions present; for example, in methyl alcohol containing $\frac{1}{4}$ per cent. of water the proportions are :—

0·0001 N.	1 ion.	500 water.	125,000 alcohol.
0·0016 N.	1 ion.	30 water.	7,500 alcohol.

If the ions had a much greater affinity for water than for methyl alcohol so that the solvent envelope consisted entirely of water molecules, there would clearly be enough water present even in the most concentrated solutions examined for the complete solvation of the ions. If the ions were solvated exclusively with water the first addition would produce a large change in the conductivity, while subsequent additions would only change the conductivity to an extent dependent upon the viscosity change. However, the effect produced by small water additions is, in every case examined, approximately proportional to the amount of water added (up to 1 per cent.) over the range of concentration studied, irrespective of whether the electrolyte is weak or strong. It is clear, therefore, that the more polarizable molecules do not monopolise the solvent atmosphere around the ions, and solvation is a statistical process, the relative amounts of the two kinds of molecules in the atmosphere being roughly proportional to the concentrations in which they are present in the solution.

This view is supported by the work of Butler[25] on the vapour-pressure of solutions of lithium chloride in alcohol-water mixtures. He finds that the activity of the water in the main bulk of the solution is reduced because the water molecules tend to congregate around the ions. Solvation, however, is not an exclusive process, for although the first 10 per cent. of water produces a larger effect than subsequent additions, further additions produce quite a considerable effect.

Effect of Water on the Degree of Association in the Alcohols.

The degree of association is given by the expression

$$\frac{\Lambda_c' - \Lambda_c}{\Lambda_c'}$$

where Λ_c is the observed conductivity, and Λ_c' is the conductivity which the salt would have if it behaved as a perfect electrolyte. For potassium chloride in methyl alcohol the

degree of association is zero. Values for other salts are given below.

Variation of Association with Water Additions.

Percentage of water.	Methyl alcohol (0·001 N.). $AgNO_3$.	Methyl alcohol (0·001 N.). NEt_4ClO_4.	Ethyl alcohol (0·0005 N.). NEt_4ClO_4.
	Percentage association.	Percentage association.	Percentage association.
None	5·2	3·6	8·9
¼		3·4	
½	4·8	3·3	8·7
1	4·3		8·5

The addition of water diminishes the degree of ion association. Before giving a reason for this decrease it is necessary to examine the effect produced by water additions on the size of the ions. An estimate of this can be made by using Stokes's equation to calculate the change in mobility due to the altered viscosity of the solvent, and comparing this with the change in Λ_0 for strong electrolytes. The percentage change in Λ_0 is less than that calculated from the viscosity, indicating that the sizes of the ions are diminished by the presence of water. However, with the exception of potassium chloride the values of Λ_0 are obtained by linear extrapolation from values of Λ_c which fall off more rapidly than they should for a perfect electrolyte and therefore give uncertain values for Λ_0. The change in Λ_c values at 0·0016 N. is also less than that predicted from the viscosity change, but here again the question is complicated by changes in the degree of association. Potassium chloride behaves as a perfect electrolyte in methyl alcohol, and in this case the depression in Λ_0 is 2 per cent. and the increase in viscosity 4 per cent., which corresponds to a decrease of 2 per cent. in the radius of the two ions. According to Ulich there is not any great difference between the degree of solvation of the potassium ion and the chloride ion, so there is no great error involved in assuming a 2 per cent. change in the radius of each ion. It will be seen in the table (Table II.) that the ions are all smaller in water than in alcohols, so it may be assumed that additions of water do actually diminish the ionic radii.

TABLE II.

Ionic Sizes in Ångström units.

Ion.	Water.	MeOH.	EtOH.	Acetone.	Lattice.
Li	2·30	3·78	5·16	3·40	0·72
Na	1·79	3·27	4·04	3·21	1·01
K	1·22	2·78	3·45	3·10	1·30
NEt_4	2·76	2·43	2·67	2·85	—
Cl	1·21	2·91	3·11	2·40	1·72
Br	1·18	2·69	2·95	2·29	1·92
I.....	1·20	2·45	2·63	2·36	2·19
ClO_4	1·33	2·11	2·23	2·23	—
Picrate	2·98	3·20	2·82	2·98	—

The above values are calculated by means of Stokes's equation. The ionic sizes in the lattice are quoted from the values of Wasastjerna [28].

The problem of ion association has been attacked by Gronwall, LaMer, and Sandved [26]. They have shown that if the potential energy of an ion in an electric field is comparable with its kinetic energy of translation, the Maxwell-Boltzmann distribution, as applied in the Debye-Hückel theory, breaks down. The effect produced is equivalent experimentally to ion association. For the purpose of the present discussion, however, the original suggestions of Bjerrum [27] afford an adequate and simpler picture. He worked out the probability of two ions being at a certain distance from one another, and found that for ions of opposite sign the probability increased very rapidly if the ions were nearer together than a certain critical distance. This distance of approach, below which the number of ion pairs, that is to say associated ions, increases rapidly is calculated to be

$$\frac{z_A z_B \epsilon^2}{2DkT}$$

where z_A and z_B are the valencies of the ions, ϵ the ionic charge, D the dielectric constant, k Boltzmann's constant, and T the absolute temperature. The values for this distance are approximately :—

Water	3·5 Å.
Methyl alcohol	9 Å.
Ethyl alcohol	11 Å.

By comparing these values with the sizes of the ions given in Table II. it will be seen that very few, and in some cases no solvent molecules can come between ions which are forming an ion pair. Even allowing for the uncertainty of the conditions in the immediate neighbourhood of the ion, and for the very approximate nature of the estimations both of the critical distance of approach and of the ionic sizes, there is still a measure of support for the view that it is largely the unsolvated ions which unite to form ion pairs. This view is rendered more probable in the light of the statistical nature of solvation, because it is likely that there are in the solution ions in all states from the unsolvated ion to the completely solvated ion. The addition of water to an alcoholic solution might be expected to increase rather than decrease the amount of Bjerrum association, because the smaller the ions the more closely they may approach, with consequent increase in the number of ion pairs. It may be that water molecules, which tend to congregate in the neighbourhood of the ions, cause a local increase in the dielectric constant. When the ions come sufficiently close together to form an ion pair practically no solvent molecules are interposed between them, and it seems likely that under the critical conditions it is the actual presence of water molecules in the solvent atmosphere which modifies the electrical forces between the ions and diminishes the probability of the occurrence of an ion pair. The effect of small water additions in reducing the degree of ion association now becomes explicable in purely qualitative terms. The reduction of the critical distance of approach from 9 Å. to 3½ Å. on passing from methyl acohol to water is by no means balanced by the reduction in ionic sizes which is proportionately less. Also, it has been shown that given equal conditions a water molecule is more likely to be found in the neighbourhood of an ion than a methyl alcohol molecule. These two factors combine to endow water with a very considerable power for reducing the degree of ion association in methyl alcohol solutions.

Results in Acetone.

The proportionality which exists between the amount of water added and the change in the conductivity is confirmed by some tests made with silver nitrate in acetone. At the end of a series of measurements with silver nitrate in acetone containing small quantities of methyl alcohol, some water was added and the change in conductivity measured ; several such mixed additions were made. The results are

given in the following table ; the first column gives the solvent added, the second column the modifying solvent already present, the third column the percentage change in conductivity, while the last column gives the percentage change when the solvent mentioned in the first column is present alone.

Solvent 1.	Solvent 2.	Change in Λ_c.	Change in Λ_c.
Water	MeOH	27	26
Water	EtOH	28	26
EtOH	MeOH	7·5	7
MeOH	EtOH	15	14

In some cases slight extrapolations have been made in the more concentrated range and the figures are liable to an error of 1 to 2 per cent., but certainly not more than 2 per cent. The agreement is remarkably close. The effects produced by equal numbers of molecules of water, methyl alcohol, and ethyl alcohol are proportional to the following numbers :—

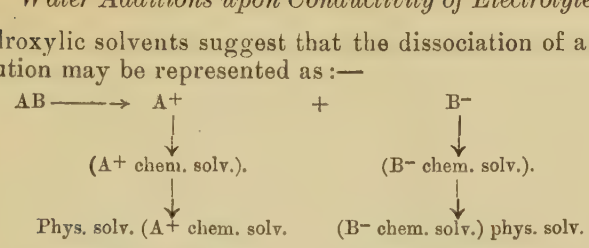
Water	16
Methyl alcohol	8
Ethyl alcohol	7

This proportionality has a further significance over and above its implication in the case of the hydroxylic solvents. In acetone there is far more ion association than in the alcohols, and water additions produce larger changes in the conductivities. This is partly explained by the low dielectric constant of acetone :—

Water	81
Methyl alcohol	30
Ethyl alcohol	25
Acetone	21

This, though a contributory cause, is not of itself sufficient. Nitromethane has a dielectric constant of 37, higher than that of methyl alcohol, but there is far more ion association in nitromethane than in methyl alcohol. Once more the key to the problem is found in ion solvation. In the hydroxylic solvents both ions can be chemically solvated, but in acetone the anion cannot be so solvated. The results obtained in the

hydroxylic solvents suggest that the dissociation of a salt in solution may be represented as :—



In acetone the stage (B^- chem. solv.) is omitted. That is a reason for the water additions having a more pronounced effect in acetone than in methyl alcohol ; over and above all questions of physical solvation there is the introduction of the stage

(B^- chem. solv. with water).

Even though the effect produced by water is large, it is surprising that it is not larger. One is forced to the conclusion that chemical solvation is not so essential to electrolytic dissociation as has been supposed. There is the point, however, that anions are not heavily solvated chemically. Whereas cations are capable of taking up four or six solvent molecules by accepting eight or twelve electrons, anions rarely solvate with more than one solvent molecule.

The change in ionic size when water is added may be calculated as before from Stokes's equation. In the alcohols there is a reduction in the size when small quantities of water are added, but in acetone there is an increase in size, with the notable exception of tetraethylammonium perchlorate. The decreases in Λ_0 are not very accurate owing to the uncertainty of the extrapolation, but the inaccuracy is not large enough to prejudice the general conclusion, namely, that the ions are larger in the presence of water. It is improbable that the introduction of the stage of chemical solvation of the anion is sufficient to account for the increase in the ionic size. There must be an increase in the degree of physical solvation of the ions. The increase in ionic size is not sufficient, of itself, to account for the decrease in the degree of association. As in the case of the alcohols, the large effect produced must be due to the presence of water in the immediate neighbourhood of the ions, and consequent modification of the properties of the medium.

Effect of Small Quantities of Water upon the Degree of Association.

It will be seen from Table I. that water additions have a small effect on the conductivities of strong electrolytes, but

their effect on weak electrolytes is considerable. If these effects are expressed, however, as changes in the percentage association of the electrolyte, as is done in Table III., the difference between strong and weak electrolytes is less marked. Small additions of water do not affect the value of the percentage association of weak electrolytes very much, although a small change in the percentage association of a weak electrolyte means a large relative increase in the degree of dissociation, and hence of the conductivity.

TABLE III.

Effect of Water upon the Degree of Association.

The values given are the percentages of salt associated at a concentration of 0·001 N.

	Methyl alcohol.			Ethyl alcohol.
Percentage of water.	$AgNO_3$.	KCl.	NEt_4ClO_4.	NEt_4ClO_4.
0	5·2	0	3·6	8·9
1	4·3	0	3·0	8·5

	Acetone.			
Percentage of water.	NEt_4ClO_4.	KI.	$HClO_4$.	KCNS.
0	7·1	5·6	11·5	13·1
1	6·2	4·3	8·1	9·5

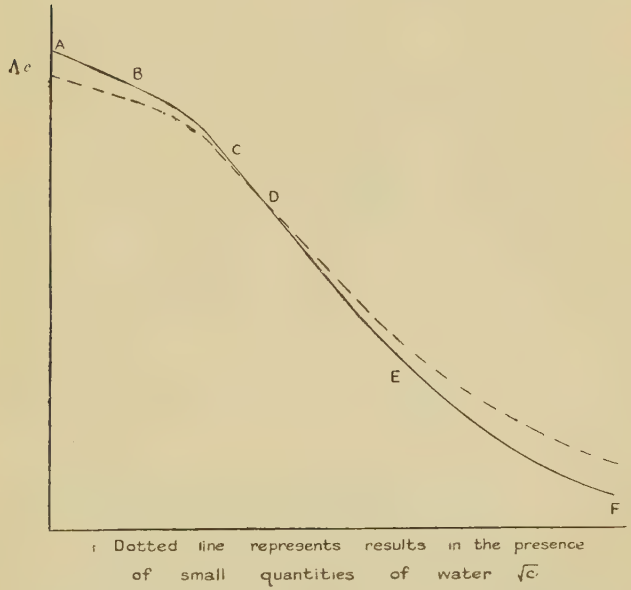
Percentage of water.	NH_4CNS.	NaCNS.	AgPic.	LiCNS.
0	44	25	34	66
1	33	18	20	46

Percentage of water.	$AgNO_3$.	Picric acid.
0	92	99
1	87	94

This raises the question as to whether the Bjerrum ion pair is of the same type as the undissociated salt, with the anion and cation united by a covalent link. To these questions the measurements made can give no definite answer, but it seems significant that there is a more or less definite series from the perfect salt (potassium chloride in methyl alcohol) through the salts with abnormally large slopes to the genuinely weak salts such as silver nitrate in acetone. The effect of water additions is growing as one passes along this series, being at its smallest for the perfect salt and increasing regularly to a maximum effect in the case

of the weakest electrolyte considered. Each electrolyte over the range of concentration employed may be regarded as being on a certain section of the general conductivity curve, and the effect of small quantities of water is given in the diagram (fig. 2). This curve is the complete curve for any electrolyte whether strong or weak, and it is because measurements are only made over a limited range that the conductivity curves seem to be different for different salts. For each salt the conductivity curve is differently placed

Fig. 2.



AB. Onsager slope; KCl in MeOH; water has little effect.
BC. NEt_4ClO_4 in MeOH and EtOH; water effect still not large.
CD. Steep but straight; water effect larger, many examples.
D. Curves intersect; AgPic in acetone and many others.
DEF. Weak electrolytes; water effect very large.

with respect to the concentration axis; all conductivity curves are essentially of the same form when complications due to solvolysis can be neglected.

Solvation and Association.

The association of ions to form ion pairs or molecules depends upon the size of the ions, and the function of

solvation is to increase the size of the ions and hence diminish the amount of association. In water most ions are large enough, even when unsolvated, to prevent sufficiently close approach for an ion pair to be formed. In solvents of lower dielectric constant solvation is necessary to make the ions large enough to resist the process of association. It is significant that the largest deviations from the Onsager slope in methyl alcohol are obtained for the tetramethylammonium salts. The tetramethylammonium ion has a large mobility in methyl alcohol, and is in consequence of its small size, very liable to associate.

Summary.

1. The effect of small water additions upon the electrical conductivity of the following solutions has been measured :—

Four salts in methyl alcohol.

One salt in ethyl alcohol.

Eight salts and two acids in acetone.

2. The viscosity of the various mixed solvents employed have been measured, and also the densities of the acetone-water mixtures.

3. The effects produced by these water additions have been found to be proportional to the amount of water added, and on this observation a statistical view of ion solvation has been put forward.

We wish to express our thanks to Mr. D. M. Murray-Rust, Mr. N. L. Ross Kane, Mr. W. A. Macfarlane, and Mr. O. Gatty for assistance in the experimental work, and to the Directors of Imperial Chemical Industries Ltd. for a grant which defrayed part of the cost of the work.

Bibliography.

(1) Debye and Hückel, *Physikal. Z.* xxiv. pp. 185. 305 (1923); Onsager, *Physikal. Z.* xxvii. p. 388 (1926); xxviii. p. 277 (1927).
(2) 'Annual Reports' (1930).
(3) Sidgwick, 'Electronic Theory of Valency' (Oxford, 1927).
(4) Schmick, *Z. Physik.* xxiv. p. 56 (1924).
(5) Washburn, J. A. C. S. xxxi. p. 322 (1909).
(6) Remy, *Z. physikal. Chem.* cxviii. p. 161 (1925); cxxiv. p. 41 (1926); Trans. Faraday Soc. xxiii. p. 381 (1927).
(7) Baborovsky, Rec. Trav. Chem. xlii. pp. 229, 533 (1923).
(8) Ulich, 'Über die Beweglichkeit der Elektrolytischen Ionen. Fortschritte der Chemie' (1926).
(9) Ulich, *Z. angew. Chem.* xli. p. 1141 (1928).

(10) Goldschmidt and Thuesen, *Z. physikal. Chem.* lxxxi. p. 30 (1912); Goldschmidt and Aarflot, *Z. physikal. Chem.* cviii. p. 121 (1924); Goldschmidt and Dahll, *Z. physikal. Chem.* cxvii. p. 312 (1925); Goldschmidt and Aarflot, *Z. physikal. Chem.* cxxii. p. 371 (1926).
(11) Murray-Rust and Hartley, Proc. Roy. Soc. A, cxxvi. p. 84 (1929).
(12) Unmack, Murray-Rust, and Hartley, Proc. Roy. Soc. A, cxxvii. p. 228 (1930).
(13) Ross Kane (unpublished).
(14) Wright, Murrey-Rust, and Hartley, J. C. S. p. 199 (1931).
(15) Hartley and Raikes, J. C. S. cxxvii. p. 524 (1925).
(16) Copley, Murray-Rust, and Hartley, J. C. S. p. 2492 (1930).
(17) Bourdillon, J. C. S. ciii. p. 791 (1913).
(18) Jones and Bingham, Amer. Chem. Jour. xxxiv. p. 481 (1906).
(19) Frazer and Hartley, Proc. Roy. Soc. A, cix. p. 351 (1925).
(20) Wynne-Jones, J. Phys. Chem. xxxi. p. 1647 (1927).
(21) Unmack, Bullock, Murray-Rust, and Hartley, Proc. Roy. Soc. A, cxxxii. p. 427 (1931).
(22) Philbrick, Dissertation (Oxford, 1925).
(23) Barak, Dissertation (Oxford, 1928).
(24) Tyndall, Proc. Roy. Soc. A, cxxix. p. 162 (1930); cxxxvi. p. 145 (1932).
(25) Butler, Proc. Roy. Soc. A, cxxix. p. 519 (1930).
(26) Gronwall, La Mer, and Sandved, *Physikal. Z.* xxix. p. 358 (1928).
(27) Bjerrum, *Kgl. Danske. Videnskabs. Selsk. Math-fys. Medd.* vii. (9) p. 1 (1926).
(28) Wasastjerna, *Soc. Scient. Fenn. Comm. Phys.-Math.* xxxviii. p. 1 (1923).

Physical Chemistry Laboratory,
Balliol College and Trinity College,
Oxford.

LII. *The Breakdown of Streamline Motion at the Higher Critical Velocity in Pipes of Circular Cross Section.* *By* Professor A. H. Gibson, *D.Sc.**

THE colour-band method of determining the higher critical velocity of flow in parallel glass tubes has been adopted by various observers since it was first devised by Osborne Reynolds. In such experiments it appears to have been the general practice to introduce the colour band at the axis of the tube.

In some preliminary experiments by the author, however, the colour-band was introduced in turn at a series of radii, with results that showed that, adopting the definition of critical velocity as the mean velocity at which eddy formation first appears at a given cross-section of the tube, the critical velocity measured in this

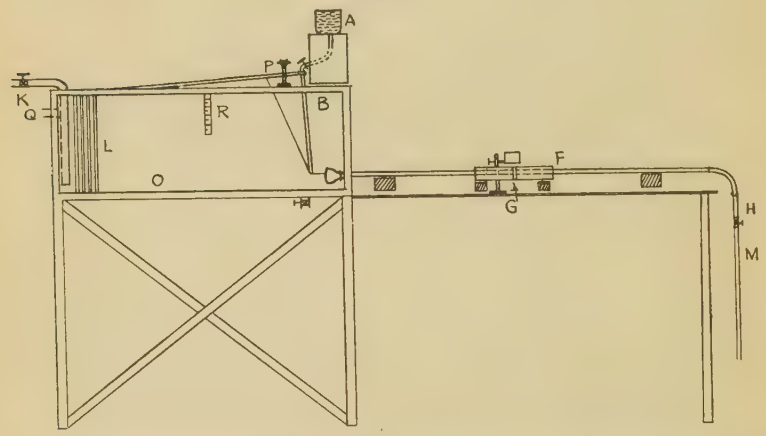
* Communicated by the Author.

way varied largely with the radial position of the coloured filament.

In order to investigate this matter further, experiments have been carried out by the author and two of his research students, Mr. A. Fogg, M.Sc., and Mr. D. L. Deshpande, M.Sc., in the Engineering Laboratory of the Manchester University, on a series of tubes of diameters ranging from 0·5 in. to 1·5 in.

The apparatus used is shown in fig. 1. It consists of a glass sided tank 5′ $10\frac{1}{2}$″ long by 18″ wide, with an inlet valve at K and baffles at L to steady the flow towards the outlet. The water from this tank discharges through the experimental glass tube, which is fitted with

Fig. 1.



a bell-mouthed entrance, the rate of flow being regulated by an outlet valve H. Aniline dye solution is supplied from the vessel A, and is conveyed to the bell-mouthed entrance of the glass tube by a copper tube terminating in a fine capillary glass tube. The copper tube is attached to the hinged platform P whose height can be regulated by an adjusting screw, enabling the position of the entering colour band to be adjusted to any required radius of the tube.

The position of the filament in the tube is observed through two horizontal telescopes mounted on adjusting screws, one above the other, in the same vertical plane. The lower telescope is used to measure the vertical distance of the filament from the axis of the tube, and the

upper one to measure its horizontal distance. For this, readings are taken in a mirror mounted above the tube at an angle of 45° to the horizontal.

In order to avoid the effect of refraction through the curved walls of the tube, that part of the tube outside the supply tank is enclosed in a water-cell G. This is 2 inches square and has vertical plate-glass sides and an open top. A graduated scale is mounted on the vertical glass front of the cell and a similar scale in the same plane across its open top.

By observing the position of the filament in each of the telescopes, relative to the corresponding scale, its horizontal and vertical distances from the axis can be determined, and its radial distance computed from the relationship $r=\sqrt{x^2+y^2}$.

In order definitely to ensure that this arrangement eliminated errors due to refraction, a thin graduated scale was fitted into a vertical diameter of the tube, and observations on this scale at different radii were compared with observations at the same radii on the outer scale. These were found to be in almost exact agreement over the whole diameter of the pipe.

Method of carrying out the Experiments.

In carrying out an experiment the supply tank is filled and allowed to stand for some time in order to allow any initial disturbance of the water to die out. Some definite cross-section of the tube at which to take observations is selected, and the telescopes and scales are brought into this plane. The outlet valve is then opened slightly and the rate of flow of the dye solution regulated so as to give a fine colour band.

The inlet valve is then opened so as to maintain a steady level in the tank, and the rate of flow through inlet and outlet is gradually increased until a slight flicker of the colour-band is noticed near the outlet end of the tube. This indicates that the velocity is approaching the critical value. A very slight further increase in the velocity causes definite but intermittent breakdown into eddy formation near the end of the tube. A further increase in the velocity causes the eddies to become permanent, and brings the point of breakdown nearer to the tube entrance. For any given radial position of the filament

there is a definite velocity at which the breakdown occurs at a given section. In carrying out an experiment the colour-band is adjusted to a pre-determined radius, and the velocity of flow is gradually increased until breakdown of the colour-band occurs at the requisite section of the tube. The experiment is then repeated with the colour-band at a different radius, until a traverse of the tube has been made.

In some of the tubes the experiments were repeated for a number of cross-sections between the entrance and exit.

The mean velocity of flow was usually determined by observing the time required for the surface level in the supply tank to fall through a measured height—usually 1 inch—when discharging with the inlet valve closed. In the experiments on the tube 1·15 in. in diameter the discharge was collected in a calibrated measuring tank. The water temperatures were also recorded at each observation.

Experimental Data.

The tubes used were of the following dimensions :—

Tube.	Diam. in.	Length. in.	Distance of point of observation from entrance to tube (in.).	Ditto expressed in pipe diameters.
a	1·50	60	24 ; 45	16 ; 30
b	1·25	51	30	24
c	1·15	72	18 ; 28 ; 42 ; 54	16 ; 24 ; 37 ; 47
d	1·0	120	60	60
e	0·5	84	19; 43 ; 72	38 ; 86 ; 144

Although, in the experiments, the motion up to the point of breakdown was always streamline, as shown by the colour-band, the inflow of water to the supply tank necessarily produced some slight initial disturbance in the water approaching the pipe. This disturbance was greater with the greater velocities of flow, and hence was greater with the larger pipes. For this reason the motion broke down with smaller values of vd/ν in the larger pipes. If v is the velocity at which breakdown occurs at the centre of the pipe at a section remote from the entrance, the values of vd/ν ranged from 8350 in the 0·5 in. pipe to 5400 in the 1·15 in. pipe. At points nearer the entrance the values of vd/ν ranged from 26,000 in the

0·5 in. pipe to 12,500 in the 1·5 in. pipe. The following table shows a typical set of observations on the pipe 1·15 in. diameter, at a section 42 inches from the entrance.

Distance of filament from centre of pipe (in.).	Temp. °F.	Mean vel. at which breakdown of filament occurs (ft. per sec.).	Corresponding value of vd/ν.
0·234	55·8	0·716	5320
0·260	56·0	0·682	5080
0·290	55·6	0·647	4820
0·299	56·3	0·629	4710
0·340	56·4	0·608	4560
0·394	56·4	0·652	4880
0·440	56·0	0·698	5210
0·450	56·0	0·716	5340
0·480	56·4	0·733	5540
0·495	56·7	0·765	5870

The curves of figs. 2–4 show the mean velocities in cm. per sec. at which breakdown occurs at each radius in the various pipes.

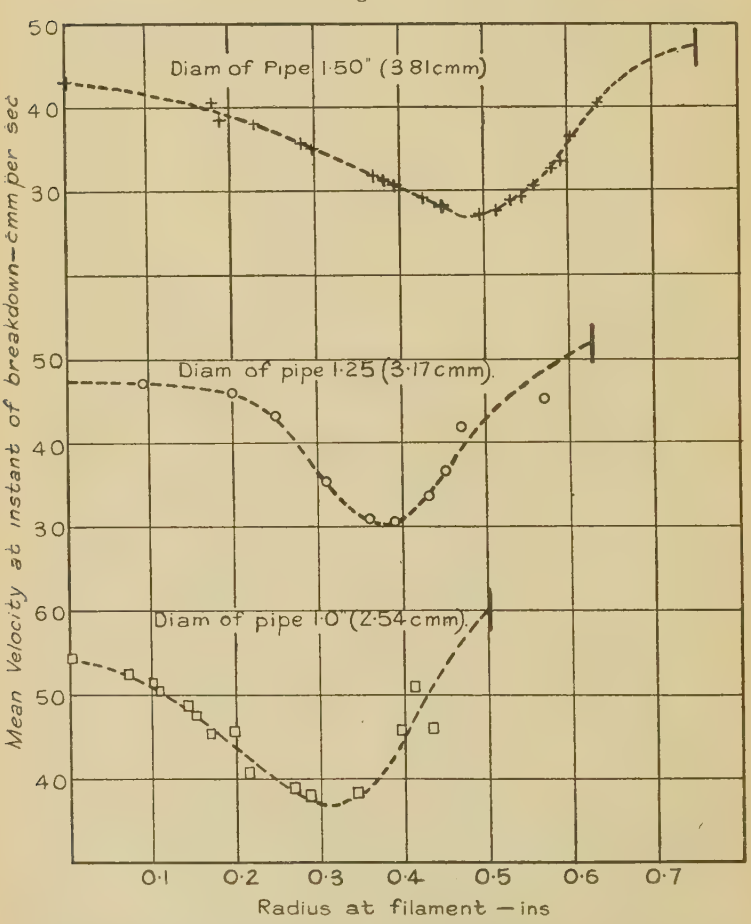
It will be seen that in each case there is a well-defined minimum. In those pipes in which measurements were made at a number of cross-sections (figs. 3 & 4) the results show that the radius at which breakdown occurs with the minimum mean velocity becomes slightly less as the distance from the bell-mouthed entrance is increased, while the difference between the mean velocity at which breaddown occurs at this critical radius and at the centre or near the wall of the pipe, diminishes as the distance from the entrance increases.

Calling r_c the critical radius and a the radius of the pipe, the values of $r_c : a$ are plotted on a base representing the distance of the section from the bell-mouthed entrance, in terms of pipe-diameters, in fig. 5, *a*. The results indicate that the ratio $r_c : a$ diminishes up to a distance of about 50 pipe-diameters from the entrance, and appears to be sensibly constant for sections more remote from the entrance. Its value diminishes somewhat as the diameter of the pipe is increased. Thus for long pipes it would appear to have a value of approximately

0·660 when the pipe-diameter is 0·5″,
0·615 ,, ,, ,, 1·0″,
0·600 ,, ,, ,, 1·5″.

If these values be plotted on a base of pipe-diameters the form of the curve would indicate that for large pipes the value is in the neighbourhood of 0·58.

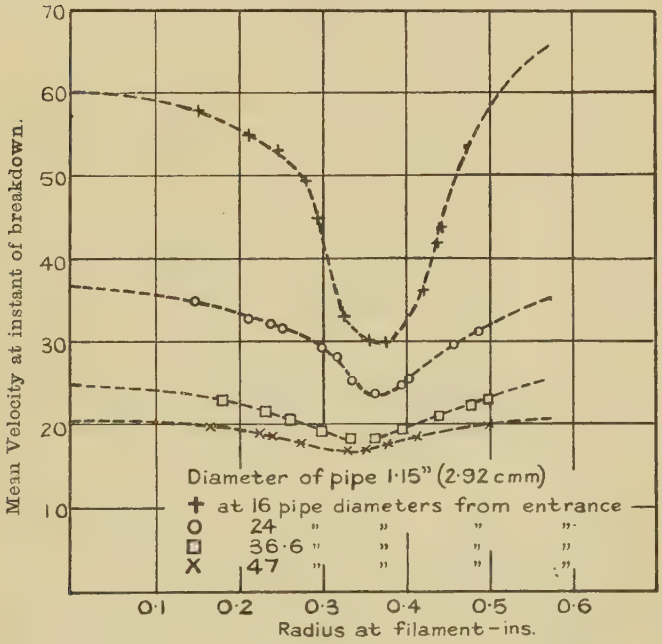
Fig. 2.



The effect of the bell-mouthed entrance is also seen in the curves of fig. 5, *b*, which show the ratio of the mean velocities necessary to cause breakdown at the centre and at the critical radius at any given section. This effect is apparent for some 80 pipe-diameters from the entrance. Beyond this point the ratio is approximately 1·10.

Owing to the difficulty of making observations very near to the walls, the form of that part of the curves of figs. 2–4 which represent the velocities near the walls, is somewhat speculative. The results, however, tend to show that the mean velocity necessary to cause breakdown to occur near the walls is somewhat higher than that necessary to cause breakdown at the centre of the pipe at the same section.

Fig. 2.

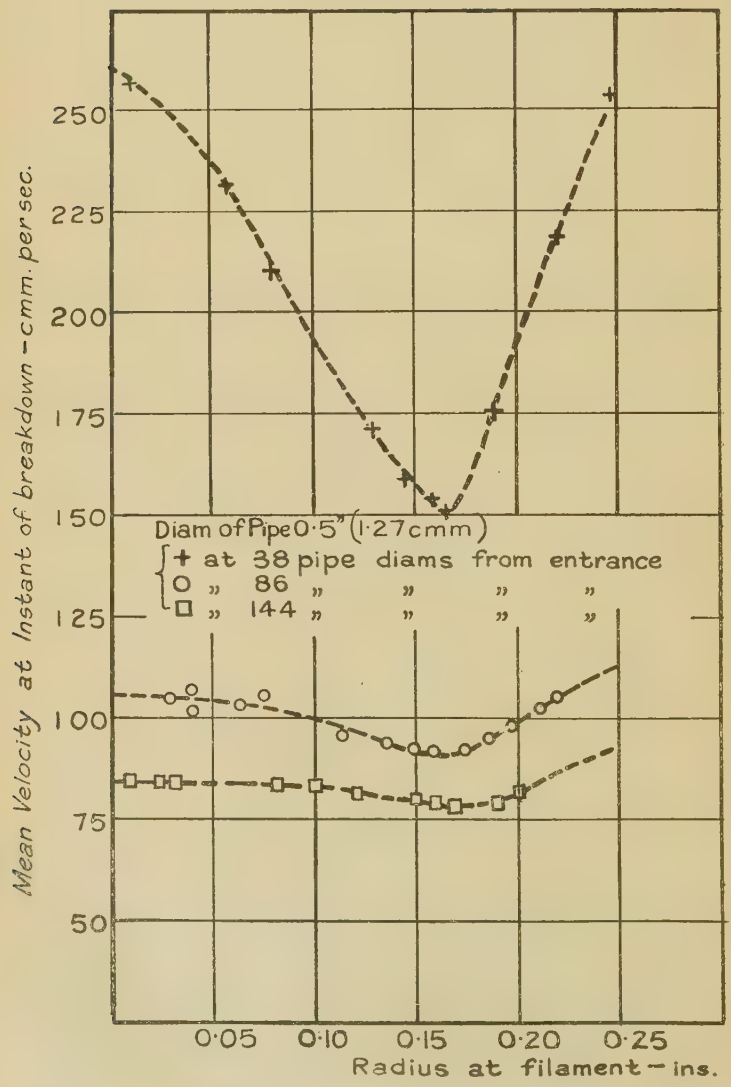


Experiments without Bell-mouthpiece.

In order to determine whether the value of the critical radius is seriously affected by the presence of the bell-mouthpiece, experiments were carried out on the 1·15 in. diameter pipe with the mouthpiece removed. In order to maintain streamline flow under these conditions, the mean velocities have to be reduced to values considerably less than those with the mouthpiece in position. Observations at a section distant 18 inches (16 pipe-diameters) from the entrance showed that under these conditions

the critical radius is 0·64 times the radius of the pipe, which is almost exactly the same as at the same section in the pipe with the mouthpiece.

Fig. 4.

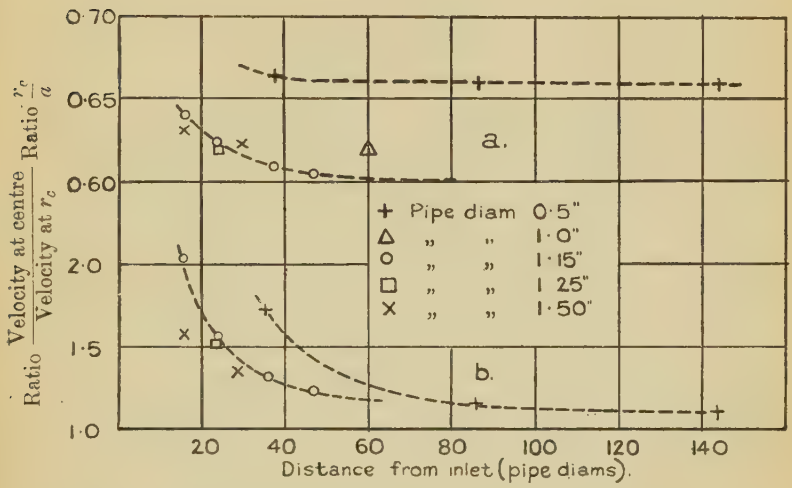


Conclusions and Deductions.

The experiments indicate that when breakdown of initially streamline motion into turbulent motion occurs during flow through a pipe, the breakdown first occurs at a distance from the centre of the pipe which is approximately 0·60 of the radius of the pipe. From this point the turbulence spreads inwards to the centre and outwards to the walls.

The value of the ratio appears to diminish somewhat as the radius of the pipe is increased, and varies from

Fig. 5.



about 0·66 in a pipe of 0·5 in. diameter to about 0·60 in a pipe 1·5 in. diameter.

The exact reason for the breakdown, or the mechanism producing it, is not yet understood. It is known to be due to the presence of the pipe walls and to be independent of the roughness of the walls so long as this roughness is small. It is not due to the attainment of a limiting shear stress in the fluid, since streamline motion is possible in a small pipe with shear stresses much greater than those existing at the critical velocity in a large pipe.

An examination of the problem in the light of the results of the present investigation suggests that the

breakdown may be related to the rate of variation of energy across a diameter of the pipe. Since the pressure across a section of a parallel tube in which the motion is truly linear and streamline is uniform, the rate of variation of the energy per unit mass along a radius is the same as the rate of change of the kinetic energy. Assuming some slight deviation of the particles from linear axial flow to be produced by some initial disturbance, this will be likely to have its maximum disturbing effect if it occurs at a radius where the radial variation of energy is a maximum, and therefore, where dv^2/dr is a maximum.

In streamline flow through a pipe of radius a,

$$v \propto (a^2 - r^2)$$

$$\therefore \frac{dv^2}{dr} \propto (a^2 - r^2)r$$

which is a maximum when $r = a/\sqrt{3} = 0{\cdot}577\ a$.

It is suggestive that this is very nearly the radius at which breakdown first occurred in the experiments. The fact that the experimental value is slightly greater than this is possibly due to the fact that in a pipe whose walls are not perfectly smooth those filaments nearest the walls suffer some slight lateral disturbance due to the roughness, and thereby suffer a reduction in stability causing breakdown to occur somewhat nearer the walls than would be the case in an ideally smooth pipe. That this is a possible explanation is indicated by the fact that primary breakdown occurred at a relatively greater radius in the smallest pipe, having a value of about $0{\cdot}66\ a$ in the smallest pipe and about $0{\cdot}60\ a$ in the largest pipe. As the surface finish was as nearly as could be determined the same in all the pipes, the relative roughness would be smallest in the largest pipes.

Viscous Flow past a Flat Surface.

If a flat surface moves tangentially through a still fluid with a velocity V, the velocity v of a particle at a distance x from the surface, and after a time of action t of the shear force is determined by

$$v = \mathrm{V} e^{-\frac{x}{a}}$$

where

$$a \propto \sqrt{\frac{ut}{\rho}}.$$

In this case $\frac{dv^2}{dx} = -\frac{V^2}{a} e^{-\frac{x}{a}}$, and is a maximum when $x=0$, *i. e.*, at the surface, and there equals $\frac{V^2}{a}$.

If, however, the surface is at rest in fluid having a velocity V,

$$v = V\left\{1 - e^{-\frac{x}{a}}\right\}$$

and

$$\frac{dv^2}{dx} = \frac{V^2}{a}\left\{1 - e^{-\frac{x}{a}}\right\} e^{-\frac{x}{a}}.$$

This is a maximum when $e^{-\frac{x}{a}} = \frac{1}{2}$, and then equals $\frac{V^2}{4a}$. The maximum value now occurs at some distance $x = 0{\cdot}69\, a$ from the surface.

It is usually tacitly assumed that the forces involved in the case of flow past a stationary body and of motion of the same body through fluid at rest, are identical; but if, as appears to be the case in flow through pipes, breakdown, due to some small initial disturbance, is most likely to occur in the region at which the lateral rate of change of kinetic energy is a maximum, there would appear to be an essential difference between them in the neighbourhood of the critical velocity.

It is to be expected, for example, that since in the case of a body moving through still fluid the breakdown is most likely to occur in the immediate vicinity of the surface, the effect of any slight roughness of the surface such as might cause initial disturbances, would be much more pronounced than in the same body if at rest in a moving fluid.

LIII. *Viscous Flow through Pipes with Cores.* *By* N. A. V. Piercy, *D.Sc.*, M. S. Hooper, *and* H. F. Winny, *Ph.D.**.

[Plate XIV.]

Introduction.

THE theoretical increase of resistance to viscous flow caused by introducing along the axis of a straight pipe a core of apparently negligible radius is well known

* Communicated by the Authors.

to be surprisingly considerable. Hitherto experiment has often appeared to conflict with this mathematical result, and doubt has arisen as to whether rectilinear flow can extensively exist in channels of wide annular section. Thus Lea and Tadros's * recent observations of unduly small resistance when the radius of the inner boundary of the channel is much less than half that of the outer boundary led them to suggest that a thin turbulent boundary layer may surround small cores and produce a semblance of slip.

In view of the apparent divergence that occurs in some circumstances between theory and experiment, further experiments were undertaken in which a pipe of circular section was fitted centrally with various cores, and it became desirable to extend the mathematical solutions for steady rectilinear flow for all forms of boundaries included in the investigation. An outstanding experimental difficulty lay in centralizing the cores with sufficient accuracy. The effect of eccentricity was investigated for this reason, and also in the hope that the solution might throw light on a swaying component observed in the motion at some speeds. This work is described in Section I. (A). Since its completion, an independent solution for the discharge has been published by Caldwell †, obtained by a different method. The brief account is given of our method on account of its comparative simplicity, and also as leading to the following extension.

A thin flat strip was also introduced as core in the experimental investigations, and to enable the results to be compared with theory it was necessary to obtain a solution for rectilinear flow in channels whose inner and outer boundaries have confocal elliptic sections. This is given in Section I. (B), and may be compared in some aspects with Lees's ‡ solution for the different case where the fluid motion results from relative axial movement of the inner and outer boundaries, and no pressure gradient exists.

The experiments form the subject of Section II.

* Phil. Mag., June 1931.
† Journ. Roy. Tech. Coll., Glasgow (1930).
‡ Proc. Roy. Soc. A, xcii. (1915).

Section I.—Theoretical.

A.—*Steady Flow through a Long Pipe with Eccentric Core of Circular Section.*

Taking the axis of z parallel to the length of the pipe, and writing P for the constant-pressure gradient, $-\frac{\partial p}{\partial z}$, and μ for the viscosity of the fluid, $\frac{\partial p}{\partial x}=\frac{\partial p}{\partial y}=0$. The remaining equation to be satisfied is

$$\nabla^2 w=-\frac{\mathrm{P}}{\mu}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

which may be arranged as

$$\nabla^2\left\{\psi-\frac{\mathrm{P}}{4\mu}(x^2+y^2)\right\}=-\frac{\mathrm{P}}{\mu}, \quad . \quad . \quad . \quad (2)$$

where ψ is a plane harmonic function making $w=0$ on the boundaries. Let the x-axis contain the centres of pipe and core, of radii a_1 and a_2 respectively, and distant b apart.

Macdonald *, studying the corresponding torsion problem, introduced the transformation $z'=c\tan\frac{\zeta}{2}$, where $z'=x+iy$ and $\zeta=\xi+i\eta$, and showed that

$$\psi_1=c^2\sum_{n=1}^{\infty}(-)^n \left\{e^{-n\beta}\coth\beta\sinh n(\eta-\alpha)-e^{-n\alpha}\coth\alpha\sinh n(\eta-\beta)\right\}\frac{}{\sinh n(\beta-\alpha)}\cos n\xi,$$

$$. \quad . \quad . \quad (3)$$

where α, β are the values of η on the outer and inner boundaries respectively, satisfies $\frac{\partial^2\psi_1}{\partial\xi^2}+\frac{\partial^2\psi_1}{\partial\eta^2}=0$, and makes $\psi_1=\frac{(x^2+y^2)}{4}+$ a constant when $\eta=\alpha$ or β. Inspection of (2) shows that

$$w=\frac{\mathrm{P}c^2}{\mu}\left\{\frac{\psi_1}{c^2}+\mathrm{A}\eta+\mathrm{B}-\frac{\cosh\eta-\cos\xi}{4(\cosh\eta+\cos\xi)}\right\}, \quad (4)$$

where A and B are constants, will satisfy (1), since the first three terms within the brackets form a plane harmonic

* Macdonald, Camb. Phil. Soc. vol. viii. (1893).

function in the ζ-plane which transforms into a plane harmonic function in the z'-plane.

On the boundaries ψ_1 is readily shown to be given by

$$\psi_1 = \frac{c^2}{2}\left(\frac{\cosh\eta}{\cosh\eta+\cos\xi} - \coth\eta\right), \quad . \quad . \quad (5)$$

so that on the boundaries

$$w = \frac{\mathrm{P}c^2}{\mu}(\mathrm{A}\eta+\mathrm{B}-\tfrac{1}{2}\coth\eta+\tfrac{1}{4}) = 0, \quad . \quad . \quad (6)$$

and, substituting α, β in turn for η, yields two equations for A and B, which together give

$$\left.\begin{aligned} \mathrm{A} &= \frac{\coth\alpha-\coth\beta}{2(\alpha-\beta)}, \\ \mathrm{B} &= \frac{\beta(1-2\coth\alpha)-\alpha(1-2\coth\beta)}{4(\alpha-\beta)}, \end{aligned}\right\} \quad . \quad . \quad . \quad (7)$$

and (4) is then completely determined.

The Flux.—The above solution is much shorter than that given by Caldwell, which makes no use of Macdonald's work and Boussinesq's well-known suggestion, but the evaluation of the volume discharge per sec.,

$$\mathrm{Q} = 2\int_0^{\pi}\int_{\alpha}^{\beta} w\left|\frac{dz'}{d\zeta}\right|^2 d\xi\, d\eta,$$

is made no simpler, and the detail sufficiently closely follows Caldwell's working for the integrations to be omitted here. Using (4), we obtain

$$\mathrm{Q} = \frac{\pi\mathrm{P}}{8\mu}\left[a_1{}^4-a_2{}^4+\frac{4b^2c^2}{\alpha-\beta}-8b^2c^2\sum_{n=1}^{\infty}\frac{ne^{-n(\alpha+\beta)}}{\sinh n(\beta-\alpha)}\right], \quad (8)$$

which agrees with Caldwell's result :

$$\mathrm{Q} = \frac{\pi\mathrm{P}}{8\mu}$$

$$\left\{a_1{}^4-a_2{}^4-\frac{(a_1+a_2+b)(a_1+a_2-b)(a_1-a_2+b)(a_1-a_2-b)}{\beta-\alpha}\right.$$

$$\left.-4b^2\left[a_2{}^2+\frac{a_2{}^4a_1{}^2}{(a_1{}^2-b^2)^2}+\frac{a_2{}^6a_1{}^4}{\{(a_1{}^2-b^2)^2-a_2{}^2b^2\}^2}+\ldots\right]\right\}$$

$$. \quad . \quad . \quad . \quad (8a)$$

Rapidly converging expansions are readily found for the summation, which also remains in the solution of the elastic problem.

The significance of (8) may most conveniently be examined by comparing with $\frac{b}{(a_1-a_2)}$ the variation of $\frac{Q}{Q_c}$, where Q_c is the discharge with the core central given by the well-known expression

$$\frac{8\mu}{\pi P}Q_c=a_1^4-a_2^4-\frac{(a_1^2-a_2^2)^2}{\log a_1/a_2}. \quad . \quad . \quad . \quad (9)$$

The results of a number of calculations with $\frac{a_2}{a_1}=0{\cdot}01$, 0·5, 0·6, and 0·7 are collected in fig. 1, where is also plotted the parabolic variation

$$\frac{Q}{Q_c}=1+\frac{3b^2}{2(a_1-a_2)^2},$$

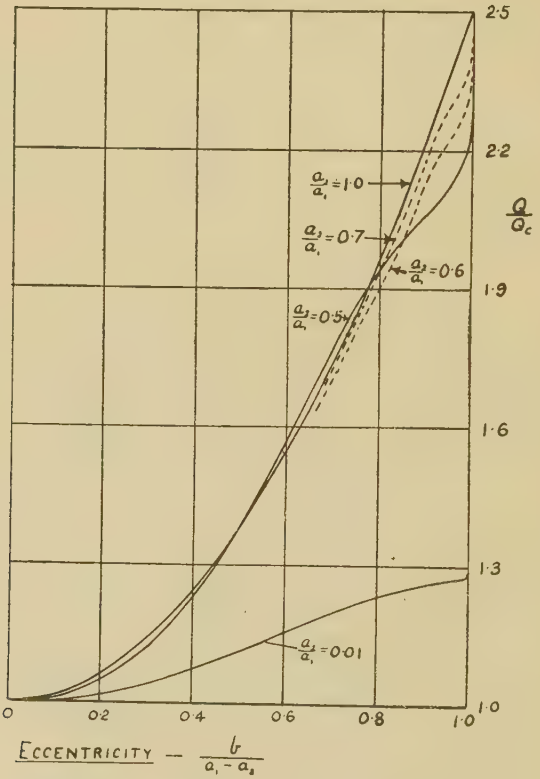
familiar * in connexion with narrow annuli. It will be seen that this simple expression, derived by assuming the flow through an element of a narrow eccentric annulus to be proportional to the cube of the distance apart of the boundaries, approximates sufficiently closely to the accurate solution for a_2/a_1 not less than 0·5 to be useful for many practical purposes in that range. With wider annuli the discharge is still considerably increased by eccentricity, but not to so great an extent. Another point of interest from (8) is that, in contrast with concentric arrangement, when a small core is at maximum eccentricity, so that it forms a small corrugation extending down the wall of the pipe, resistance is scarcely affected.

Distribution of Velocity.—Iso-velocity lines have been prepared for the case $\frac{a_2}{a_1}=\frac{b}{a_1}=0{\cdot}25$, and are shown for equal intervals of velocity in fig. 2. Inspection readily illustrates a marked falling away of shear on the restricted side of the channel. Between the internal and external boundaries where they most closely approach one another the maximum velocity is less than one-third

* The large increase of leakage, for instance, through a nearly closed circular aperture when the plug binds to one side has long been recognized by engineers.

that on the opposite side, although the flux is increased only by some 12 per cent. As eccentricity is decreased a difference of this kind in the velocity distribution tends to remain of significance longer than the reduction of resistance. The reflexion that small eccentricities may

Fig. 1.



Variation of discharge with eccentricity of core.

be expected to exist in any experimental annulus, often in wave form, owing to the mechanical difficulty of locating a core centrally within a pipe with accuracy except at isolated positions along the length, suggests from the above that in streamline motion a disturbance of swaying type may commonly arise.

Resistance of Pipe and of Core.—Let F_{pc} and F_c be the tractions per unit length along the outer and inner boundaries respectively. In the z'-plane

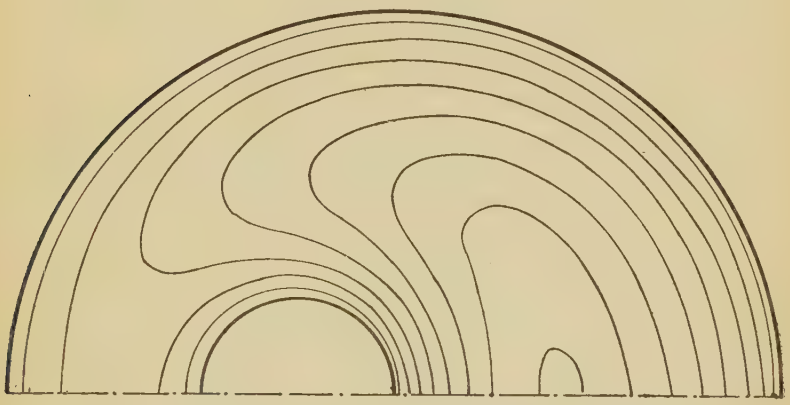
$$F=\mu\int\left(\frac{\partial w}{\partial n}\right)_{n=0} ds,$$

and transforming we have for the pipe

$$F_{pc}=2\mu\int_0^{\pi}\left(\frac{\partial w}{\partial \eta}\right)_{\eta=a} d\xi. \quad . \quad . \quad . \quad . \quad (10)$$

Applying (10) to the terms of (4) taken in order, the first

Fig. 2.



Iso-velocity lines for circular pipes with eccentric cores.

and third terms vanish, and the second and fourth easily reduce to give

$$F_{pc}=\pi P\left(a_1^2+\frac{bc}{a-\beta}\right). \quad . \quad . \quad . \quad . \quad (11)$$

Similarly the force per unit length along the core is found to be

$$F_c=-\pi P\left(a_2^2+\frac{bc}{a-\beta}\right). \quad . \quad . \quad . \quad . \quad (12)$$

These expressions show that the mean intensities of friction along core and pipe are little changed until eccentricity becomes appreciable, but tend towards equality as the core approaches the wall of the pipe.

B.—*Steady Flow through a Long Pipe of Elliptic Section with a Confocal Elliptic Core.*

Using the transformation

$$z' = c \cosh \zeta, \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

i. e., $x = c \cos \eta \cosh \xi$ and $y = c \sin \eta \sinh \xi$, and proceeding as in Section I. for eccentric circular cores, it is known * that

$$\psi_1 = -\frac{c^2}{8} \cdot \frac{\sinh 2(\alpha-\xi)+\sinh 2(\xi-\beta)}{\sinh 2(\alpha-\beta)} \cos 2\eta, \quad (14)$$

which satisfies $\frac{\partial^2 \psi_1}{\partial \xi^2} + \frac{\partial^2 \psi_1}{\partial \eta^2} = 0$, would make w constant on the boundaries. Hence the velocity is determined by

$$w = -\frac{Pc^2}{\mu} \left\{ \frac{\psi_1}{c^2} + A\xi + B + \frac{\cosh^2 \xi - \sin^2 \eta}{4} \right\}, \quad (15)$$

when the constants A and B have been found from the condition $w = 0$ on the boundaries. Putting $\xi = \alpha, \beta$ on the outer and inner boundaries respectively, we obtain

$$\left.\begin{aligned} A &= \frac{(\cosh^2 \beta - \cosh^2 \alpha)}{4(\alpha-\beta)}, \\ B &= \frac{(\beta \cosh 2\alpha - \alpha \cosh 2\beta)}{8(\alpha-\beta)}. \end{aligned}\right\} \quad . \quad . \quad . \quad (16)$$

The Flux.—Q, the volume passing per sec. through half the section, is given by

$$Q = \int_\beta^\alpha \int_0^\pi wc^2 (\sinh^2 \xi + \sin^2 \eta)\, d\xi\, d\eta, \quad . \quad . \quad (17)$$

and denoting by Q_1, Q_2, etc., the contributions to Q by the terms within the brackets of (15) taken in order,

$$\begin{aligned} Q_1 &= \int_\beta^\alpha \int_0^\pi \psi_1 (\sinh^2 \xi + \sinh^2 \eta) d\xi\, d\eta \\ &= -\frac{c^2 \pi}{32} \cdot \frac{1-\cosh 2(\alpha-\beta)}{\sinh (2\alpha-\beta)}; \end{aligned}$$

$$Q_2 = \frac{c^2 \pi A}{4} \left[\xi \sinh 2\xi - \frac{\cosh 2\xi}{2} \right]_\beta^\alpha.$$

* Love, ' Elasticity,' 2nd ed., p. 308

$$Q_3 = \frac{c^2\pi B}{4}[\sinh 2\xi]_\beta^\alpha,$$

$$Q_4 = \frac{c^2\pi}{16}\left[\frac{\sinh 4\xi}{8}\right]_\beta^\alpha.$$

The total flux then reduces to

$$2Q = \frac{\pi P c^4}{64\mu}\left\{\sinh 4\alpha - \sinh 4\beta - 4\tanh(\alpha-\beta) - \frac{2(\cosh 2\alpha - \cosh 2\beta)^2}{\alpha-\beta}\right\}. \quad (18)$$

Resistance.—The traction per unit length along the internal boundary is

$$F_c = -2Pc^2\int_0^\pi \frac{\partial}{\partial\xi}\left\{\frac{\psi_1}{c^2} + A\xi + B + \frac{\cosh^2\xi - \sin^2\eta}{4}\right\}_{\xi=\beta} . d\eta. \quad . \quad . \quad . \quad (19)$$

On integration the first and third terms within the brackets of (19) vanish, so that

$$F_c = -2\pi Pc^2\left(\frac{\sinh 2\beta}{4} + A\right). \quad . \quad . \quad . \quad (20)$$

Similarly, for the force on the external boundary,

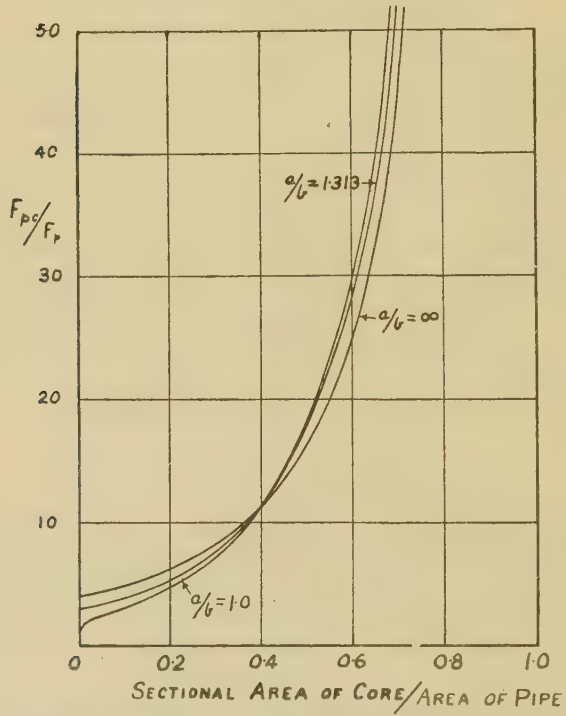
$$F_{pc} = 2\pi Pc^2\left(\frac{\sinh 2\alpha}{4} + A\right). \quad . \quad . \quad . \quad (21)$$

Variation of Pipe Friction with Size of Core.—From the foregoing we have calculated for constant discharge the ratio of the traction per unit length along the pipe surface when a central core is present to the same when no core exists, *i. e.*, the ratio $\frac{F_{pc}}{F_p}$, through a wide range of relative sizes of pipe and core. If a denotes the major and b here the minor semi-axis of the pipe section, the calculations relate to the three values of $\frac{a}{b}$: 1, 1·313, and ∞. The results are illustrated in fig. 3, where the three curves give the variation of $\frac{F_{pc}}{F_p}$ against that of the ratio of the sectional areas of core and pipe for the above three values of $\frac{a}{b}$. It will be seen that when

the sectional area of the core exceeds one-third that of the pipe, the proportionate increase of traction, due to the introduction of the core, along the outer boundaries of channels formed between confocal elliptic walls approximates, over a considerable range of mean width of channel, to that for channels of annular section.

A Flat Strip as Core.—An interesting case occurs when

Fig. 3.

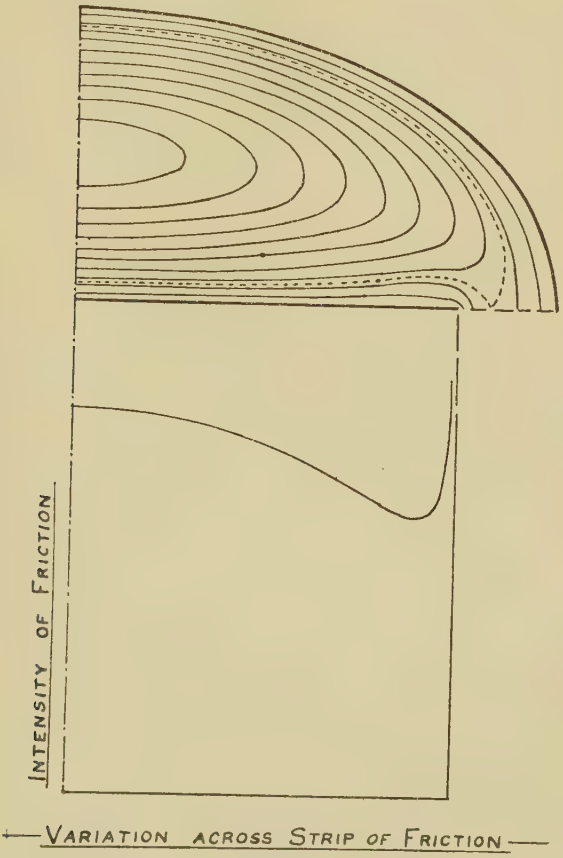


Variation of pipe-friction with size of core—elliptic pipe and core.

the inner boundary contracts to the straight line joining the foci of the outer boundary, so that $\xi = \beta = 0$. Iso-velocity lines have been prepared for this case with $\frac{b}{a} = 0{\cdot}6$, and are shown in fig. 4, where the increments of velocity are equal, excepting the two critical contours, drawn as broken lines, which cut in the plane of the strip.

It can be made out from this figure that the intensity of the force along the outer boundary is least in the region of greatest curvature. Along the strip it first decreases from its central value before passing to ∞ at the edges.

Fig. 4.



Iso-velocity lines for elliptic pipe with strip as core.

The ratio of the intensity of friction at the position η to that at the centre of the strip is readily found to be

$$\frac{\operatorname{cosec}\eta(\sinh 2\alpha - 2\alpha\cos 2\eta)}{(\sinh 2\alpha + 2\alpha)}.$$

Evaluating this ratio in the present case gave the result also shown in fig. 4.

Lees *, studying the resistance of cores moving slowly through viscous fluid contained in stationary pipes, obtained the interesting result for this motion that the resistance of a thin flat strip moving in a wide confocal pipe is the same as that of a circular cylinder of diameter equal to half its width, moving in a circular pipe of radius equal to the mean of major and minor semi-axes of the elliptical pipe. In the present case of motion this result can be shown to hold only when the width of the strip becomes very small compared with the mean diameter of the pipe ; otherwise the difference in mean intensity of friction is less marked.

Section II.—The Experiments.

Earlier Experimental Investigations.—Most published accounts of experiments on flow through cored pipes have been concerned chiefly with narrow annuli and with examining the feasibility of a simple correlation of resistance in eddying flow on the basis of "hydraulic mean depth." Thus Becker † used only narrow annuli for his measurements of the flow of water, air, and steam. Atherton's ‡ experiments with water, air and oil covered a wider range, the ratios of the inner to the outer diameters of the channels employed being 0·36, 0·45, and 0·80, but his results are in conflict with other measurements, and their present interest is obscured by the provision of a "stilling length" of only 15 diameters for disturbances caused by the front ends of the cores. Caldwell's § experiments return to the large diameter ratio of 0·844, and their particular interest lies in verifying for narrow annuli the calculated effect on resistance of eccentricity.

The most relevant published investigation is that of Lea and Tadros ||, which has already been mentioned. Their pipe was horizontal, as usual, but whilst previous arrangements employed faired distance-pins to support the core from the pipe-wall, in the present instance such observations were omitted, the wires forming the

* *Loc. cit.*
† *Ver. Deut. Ing. Zeit.* vol. li. p. 1133 (1907).
‡ Amer. Soc. Mech. E. Trans. vol. xlviii. (1926).
§ *Loc. cit.*
|| *Loc. cit.*

cores being passed over a pulley and loaded to minimize sagging. At low Reynolds numbers experiment agreed with theory with a core-to-pipe diameter ratio of 0·5, but when this was reduced to 0·185, for example, the discharge observed exceeded the theoretical by 70 per cent., and for still smaller cores the discrepancy became much more marked. As a consequence it was suggested that turbulent velocities at entry failed to be damped out adjacent to the convex surface, leaving a boundary layer which simulated slip in its effect on resistance. An early first critical speed and a slow transition to general turbulence were recorded.

The analysis of Section I. of the present paper rules out a possibility of explaining the above gross differences by accidental eccentricities. But it does not follow from the experiments or the explanation advanced that damping of inlet disturbances is zero or negative near the inner boundary of a wide annular channel, because the question is left unanswered as to whether a greater stilling length would not serve to develop the strictly rectilinear velocity profile at experimental sections farther downstream; the case might be similar in some ways to that shown by Schiller * to hold along the inlet length of a pipe.

So far as we have discovered, no experiments have previously been carried out with a strip as core or with a channel bounded by confocal elliptic surfaces.

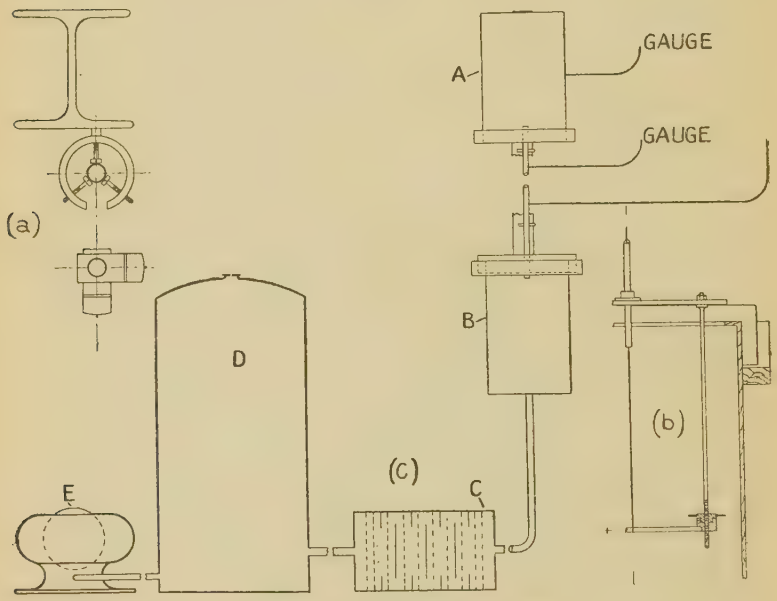
Friction Measurements, Apparatus, and Method.—The present investigation deals for the most part with measurements of frictional resistance, which will now be considered. Later an account will be given of some visual observations.

Preliminary experiments gave results somewhat resembling those of Lea and Tadros, but discrepancies were traced to curvature of the pipe and variable eccentricity of the core. It was then decided to employ air as fluid flowing down a vertical pipe, for which use was made of the head-room of a small theatre at East London College. Special precautions were taken to constrain the pipe to straightness with some accuracy. The cores were made laterally adjustable, but were freed from any form of lateral support within the pipe. It was also provided to deal with exceptionally small rates of discharge if necessary.

* *Zeit. f. a. Math. u. Mech.* vol. ii. (1922).

The vertically arranged smooth-drawn brass pipe, of 1·288 cm. diameter and 360 diameters length, was selected with some success for straightness, and was elastically strained to a closer approximation by being clamped within adjustable collars spaced at intervals of 60 cm. along a rigid iron girder. Near each clamp two sectors, accurately turned to a constant radius, were fixed to the pipe mutually at right angles (fig. 5, *a*), and

Fig. 5.



Arrangement of apparatus.

the position of the pipe within the collars was adjusted until one of each pair of sectors just touched one or other of two fine straightened wires tautly strung parallel to the pipe. Contact was observed electrically, the sectors being insulated. The pipe was not fitted with a bell-mouth entry, and no special precautions were taken to ensure absence to any complete degree of turbulence in the fluid at entry ; the flow through a central length without a core in place changed from streamline to turbulent form at a Reynolds number in close agreement with Stanton and Panell's investigations and others.

Near each end the pipe passed through a stout strip carrying two columns (fig. 5, *b*), supporting between them a narrow adjustable beam of faired section from which the cores were suspended, their ends projecting well beyond the pipe. A fitment was added to the lower pair of columns by which considerable tension could be applied to the cores.

The smaller cores, of brass or copper wire, were set to straightness under tension and heat, being carefully measured, after polishing, to verify absence of any appreciable deformation. The limited loading available in the pipe made this process inadequate as diameter increased, and for the larger cores high-tension electric cables were used ; they proved sufficiently flexible to make the limited tension effective after previous straightening, and it was found possible to select closely circular and uniform lengths. Their surface was polished smooth with graphite. The mean diameters of the eleven cores of circular section were : 0·0022, 0·0075, 0·0200, 0·0260, 0·0400, 0·0911, 0·1675, 0·236, 0·373, 0·635, 0·787 cm. Details of the flat strip also used as core are given later.

The air-circuit is indicated at *c* (fig. 5). The upper end of the pipe projected about 14 cm. into a large sealed box (A) provided with an orifice for measuring intake, while a second box (B) was similarly situated at the lower end. Air was drawn by way of these boxes through a third (C), obstructed by wire-gauze baffles, to a 900-litre low-pressure reservoir (D), which was maintained partially exhausted by an evacuator (E) driven from batteries. The evacuator was run continuously at full speed, velocity through the pipe being changed by means of an adjustable leak of the external air into the low-pressure tank. Steady flow resulted, except in windy weather. Certain details of the arrangement were adopted from an investigation by Watson and Schofield *.

One or other of two sharp-edged orifices, of diameters 0·475 and 0·238 cm., was employed to measure the flow. They were calibrated in place by an aspirator method. While the larger one showed no appreciable variation from the familiar approximate law of flow at considerable speeds through orifices, the smaller showed viscous "scale effect," which was taken duly into account.

* Proc. Inst. Mech. E., May 1912.

The resistance of each annulus was assessed at about 40 speeds, ranging from 0·021 to 2·7 times the Reynolds critical speed for the pipe alone, each estimate being the mean of five to ten readings taken at very nearly constant speed. The pressure-differences for this purpose were observed between two small holes in the pipe, alternatively 220 and 130 pipe-diameters apart. Some comparisons made between the alternative lengths with one of the smaller cores showed that differences could not be made good against accidental variations and, for accuracy in the measurements, the more widely spaced pair of pressure-holes was chosen. The upstream pressure-point was then 60 pipe-diameters away from the inlet. The Chattock gauges employed were specially adapted to cope with the unusually small pressure-differences to be measured in some cases, and suitable precautions were taken to correct their readings for temperature effects.

It is convenient to introduce v for the mean velocity through the annulus and D for $2a_1$ the diameter of the pipe. Rayleigh's well known general formula for resistance of the present kind with any one annulus may be written

$$\frac{\mathrm{F}}{\rho v^2 \pi \mathrm{D}} = f\left(\frac{v\mathrm{D}}{\nu}\right),$$

and a special case that proves useful in examining the present result is

$$\frac{\mathrm{F}}{\rho v^2 \pi \mathrm{D}} = k\left(\frac{\nu}{v\mathrm{D}}\right)^s, \quad . \quad . \quad . \quad . \quad . \quad (22)$$

where s is a numerical index and k a constant. When s has the value unity, (22) becomes, for rectilinear streamline flow,

$$\frac{\mathrm{F}}{\mu v \pi} = k'. \quad . \quad . \quad . \quad . \quad . \quad . \quad (23)$$

Realization of constant values for k' indicates rectilinear flow through the range of v, and all measurements through this range are conveniently collected and expressed in a mean value of k' for each annulus. This might form the basis of a comparison with theory, but such comparison is more conveniently provided for as follows.

The mathematical solution for concentric cores of circular section may be written

$$\left.\begin{aligned} F_c &= 4\pi\mu v\{f(a_2/a_1)-1\}, \\ F_{pc} &= 4\pi\mu v\{f(a_2/a_1)+1\}, \\ \text{where} \\ f(a_2/a_1) &= 1/\{(a_1^2+a_2^2)/a_1^2-a_2^2)-1/\log(a_1/a_2\}, \end{aligned}\right\} \quad (24)$$

and F_c denotes, as before, the friction per unit length of the core and F_{pc} that of the pipe-wall when the core is in place. Comparison of (24) with the friction per unit length, $F_p = 8\pi\mu v$, for a pipe without a core shows that

$$\frac{(F_c+F_{pc})}{F_p} = f\left(\frac{a_2}{a_1}\right). \quad . \quad . \quad . \quad . \quad . \quad . \quad (25)$$

The ratio of the L.H.S. is directly obtainable from the readings, and may be compared with calculated values of $f\left(\frac{a_2}{a_1}\right)$.

In assessing critical speeds the well known systematic method suggested by (22) was followed. It was found, on plotting $\log F$ against $\log\left(\frac{vD}{\nu}\right)$, that all the observations for each annulus lay within small experimental errors along a number of straight lines until well into turbulent flow, and the critical speeds were defined by the intersections of these lines.

The Friction Results.—These are summarized in Table I. The first column gives the radii of the cores expressed in terms of the radius of the pipe. The second column gives the mean experimental value of k' as far as the first critical speed ; it will be seen that without a core in place this quantity differs by only 0·6 per cent. from the corresponding mathematical result, and a satisfactory check is thus provided on various calibrations entering into the work. In the fourth column are given the ratios of the mean values of the resistance of the annuli to that of the pipe alone, and these may be compared with the calculated values of $f\left(\frac{a_2}{a_1}\right)$ which are collected in the third column. In fig. 6 the curve (1) is calculated, while the points are experimental values.

Below the first critical speed it will be noticed that with a diameter-ratio of 0·0202 the mathematical result for rectilinear flow is realized within 0·8 per cent., and the theoretical difference due to the core to within 3 per cent. All readings taken are included in the means given in the table ; with lower readings rejected, experiment would have agreed with theory, beyond the above diameter-ratio and ignoring that of 0·13, within 2 per cent.

TABLE I.

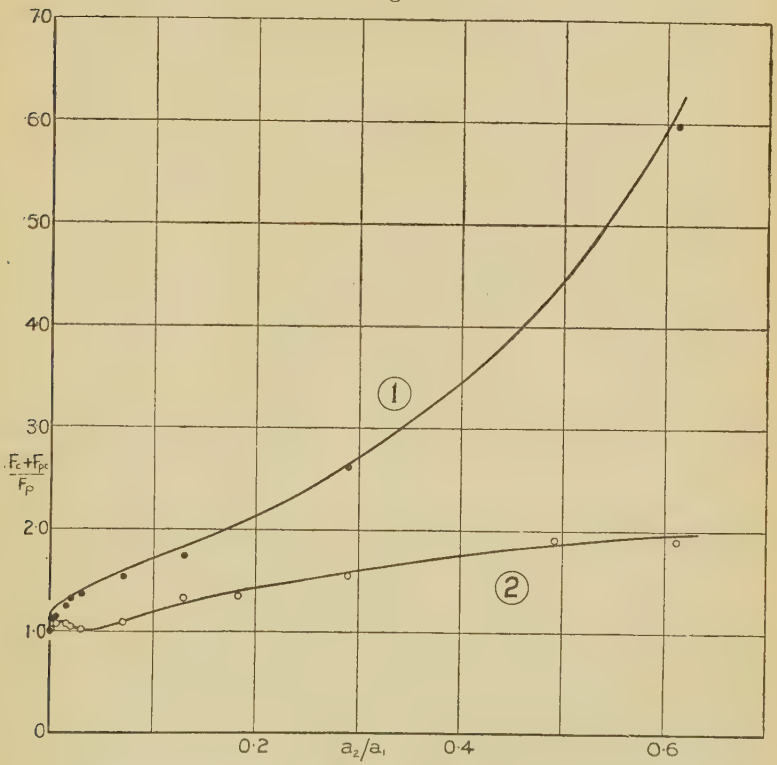
(1).	(2).	(3). Streamline flow.		(4).	(5). Turbulent flow.	(6).
$\frac{a_2}{a_1}$.	$\pi k'$.	$\frac{(F_c+F_{pc})}{F_p}$.			$\frac{F}{\rho v^2 \pi D}$.	$\frac{(F_c+F_{pc})}{F_p}$.
		Theory.	Experiment.			
Pipe	24·98	1·000	1·000		0·00570	1·000
0·0017	27·9	1·186	1·117		0·00578	1·106
0·0058	28·7	1·241	1·149		0·00607	1·066
0·0155	31·1	1·315	1·245		0·00609	1·068
0·0202	33·3	1·343	1·333		0·00595	1·042
0·0311	34·1	1·401	1·365		0·00578	1·017
0·0707	38·1	1·581	1·525		0·00620	1·086
0·130	43·2	1·837	1·729		0·00755	1·326
0·183	—	—	—		0·00765	1·343
0·290	65·2	2·659	2·610		0·0088	1·543
0·493	—	—	—		0·0109	1·915
0·612	149·0	6·190	5·965		0·0105	1·91

Thus with the present apparatus any divergence between theory and experiment, except with very small cores, must be due to eccentricity and errors of measurement, a result in sharp disagreement with the findings of Lea and Tadros.

Discrepancies are greater with the three small cores; even rejecting lower readings these vary between 4 per cent. and 60 per cent. Further experiments within the range continued to show unaccounted differences from theory which changed with resettings of the core

to a degree unexplainable by uniform eccentricity, according to Section I. It was concluded that with a diameter-ratio less than 0·01 rectilinear flow could not be established with the present apparatus, as Lea and Tadros found so much more extensively in their case. The present success obtained with double this lower

Fig. 6.



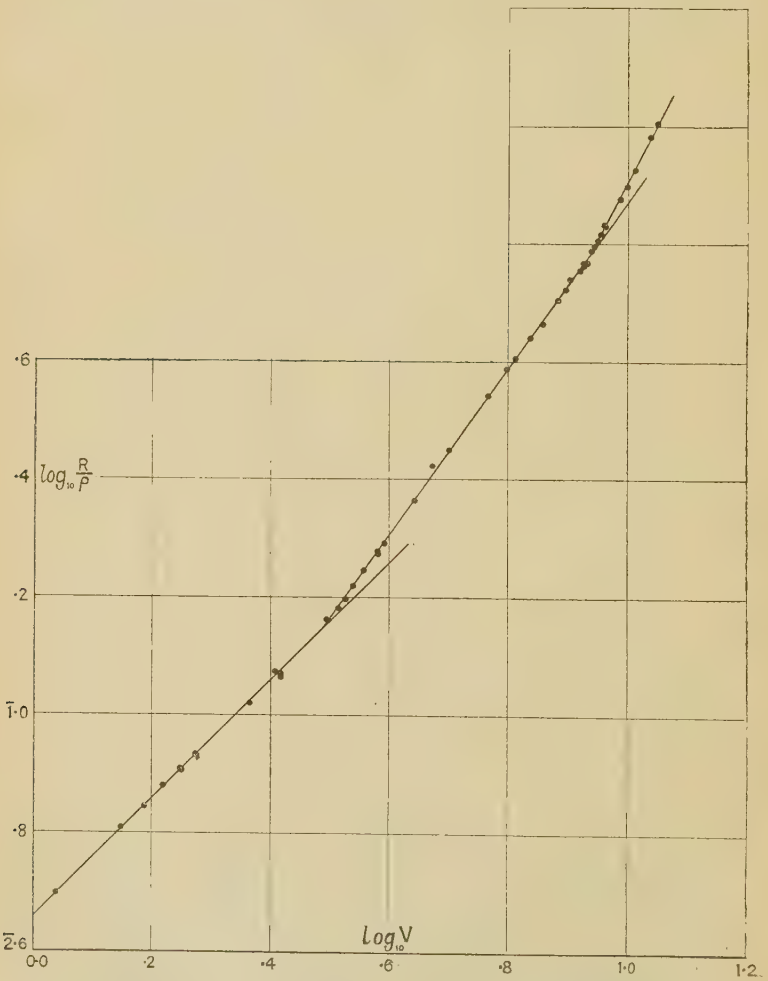
(1) Parallel streamline flow.—Curve theoretical, points experimental.
(2) Turbulent flow.

limit suggests, however, that apparatus with a longer stilling length or considerably reduced inlet disturbances, and which further eliminated variation of eccentricity along the channel, would establish rectilinear flow with cores of still smaller size.

The last two columns of Table I. refer to motion with fully developed turbulence. To collect the readings

it has been decided to give in column 5 mean values of $\frac{F}{\rho v^2 \pi D}$ for the range of $\frac{vD}{\nu}$: 2900–3400 (but always

Fig. 7.



A typical set of observations, determining critical speeds.

beyond the last critical speed), except for the diameter-ratio 0·612, for which the higher range, 5300–5580, has required to be substituted. These mean results give

curve (2) of fig. 6. The increase of resistance due to the introduction of a core concentrically within a pipe is seen to be notably less in turbulent than in streamline flow; a rough gauge obtainable from the present experiments is that it is four times as important in streamline flow. The increase loses significance in a progressive manner through the wide range existing between steady flow and fully developed turbulence.

The Critical Speeds.—Fig. 7 illustrates by a typical example the definition obtainable in estimating critical values of $v\mathrm{D}/\nu$ by the method already described.

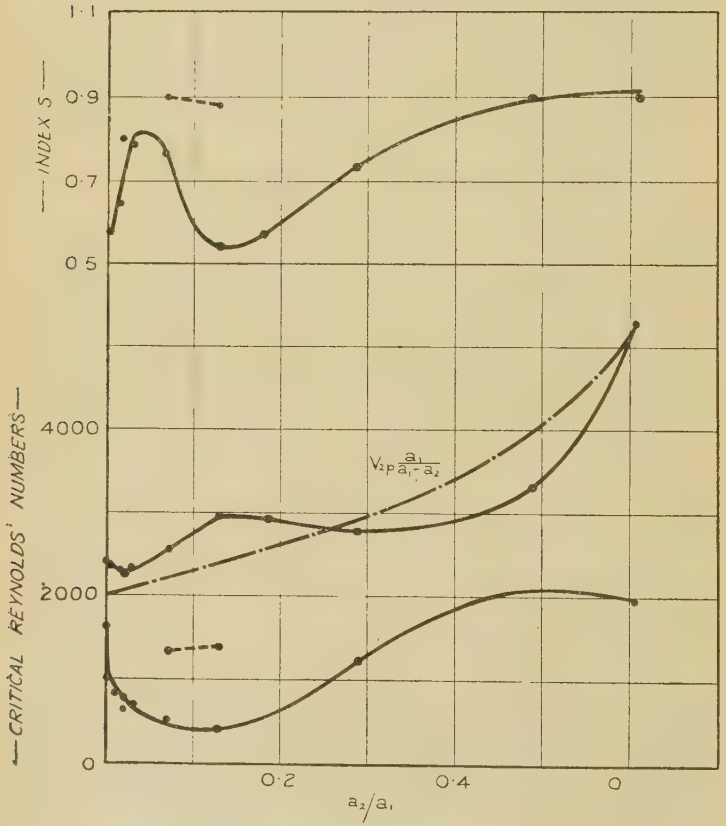
TABLE II.

(1). $\frac{a_2}{a_1}$.	(2). A.	(2 *a*). $\frac{k_1}{2}$.	(3). *s*.	(4). C.	(4 *a*). $\frac{k_2}{2}$.	(5). k_D.
Pipe	1620	—	0·90	2060	2060	0·0041
0·0017	1022	80·1	0·55	2380	2190	0·0056
0·0058	825	80·1	0·59	2340	2110	0·0061
0·0155	770	92·6	0·65	2310	2030	0·0062
0·0202	653	83·4	0·80	2240	1960	0·0062
0·0311	690	98·8	0·78	2340	2010	0·0060
0·0707	545	100·2	0·90	2520	2060	0·0058
0·130	400	91·1	0·88	2950	2270	0·0068
0·183	—	—	0·58	2920	2180	0·0066
0·290	1250	(390)	0·73	2760	1900	0·0074
0·493	—	—	0·90	3320	2010	0·0072
0·612	1910	(804)	0·90	5300	(3070)	0·0062

Referring to Table II., the limiting Reynolds number (A) for steady flow is given for the several annuli in column 2, while the Reynolds number (C) which marks the supervention of general turbulence is given in column 4. With two cores there occur intermediate critical speeds (B)—1330 and 1380 for $\frac{a_2}{a_1} = 0{\cdot}0707$ and 0·130 respectively. Between the critical speeds (A) and (C) or (A) and (B) the index *s* of (22) has, within small experimental errors, the constant values given in column 3. Between (B) and (C) for the smaller and larger of the exceptional cores, respectively, $s = 0{\cdot}76$ and 0·55. The results are plotted in fig. 8.

It will be seen that the introduction of a very small core reduces by some 40 per cent. the range for streamline flow in a pipe under the present conditions, and delays general turbulence by some 15 per cent. The observed small preliminary decrease of s for a pipe

Fig. 8.



Critical speeds for parallel streamline flow and fully developed turbulence,
Chain-line: hydraulic mean depth approximation.
Above: index of friction law between critical speeds.

without a core agrees with other experiments, and points to the stilling length being on the short side. The reduced width of the annulus begins to lead to a high second critical speed when the radius of the core attains

half that of the pipe. But there continues to exist a low critical speed, so that it seems probable that flow through a narrow annulus, unless very long, would be associated over a wide range with a value of s less than unity.

Correlation of Critical Speeds.—The measurements of columns 2 and 4 of Table II., prove suggestive on further examination. If v_1 denotes the first and v_2 the second critical speed, the observations for wide annuli may be expressed approximately in the form

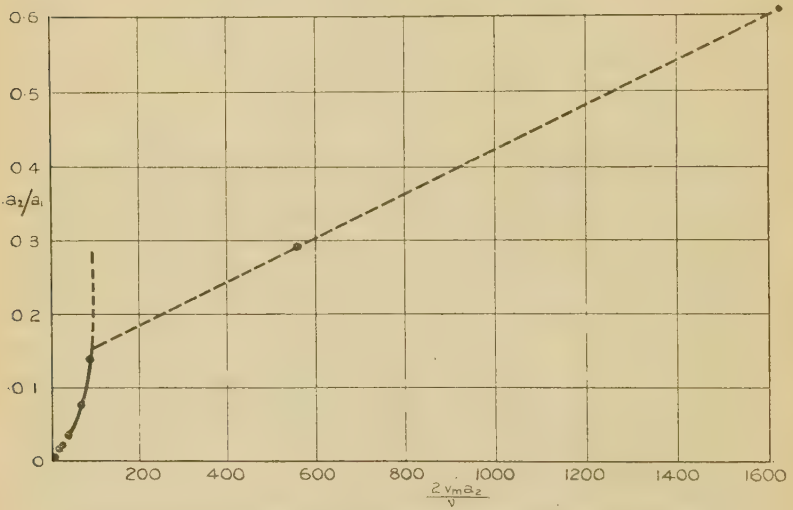
$$\left.\begin{aligned} \frac{v_1 D}{\nu} &= k_1 \frac{f(a_2/a_1)}{f(a_2/a_1)-1}, \\ \frac{v_2 D}{\nu} &= k_2 \frac{f(a_2/a_1)}{f(a_2/a_1)+1}. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (26)$$

Columns $2a$ and $4a$ of Table II. indicate the measure of success of these formulæ ; for v_1 an extreme variation from the mean of 46 per cent. is reduced to 12 per cent., while for v_2 a maximum variation of 30 per cent. from the mean is reduced to 8 per cent. The mean value of k_2 is within 1 per cent. of its value for the pipe alone. Comparison of (26) and (25) shows that, to the above approximation, and for wide annuli only, v_1 is inversely proportional to the ratio of the force exerted by the core on the fluid to the whole force, while if the ratio of the forces on the two boundaries just prior to the occurrence of v_2 be assumed the same as for rectilinear flow, v_2 is inversely proportional to the ratio of the force exerted by the pipe-wall on the fluid to the whole force.

The foregoing suggests the possibility of ascribing the first critical speed to the core, and the second mainly to the pipe, for wide annuli. When the core is very small compared with the pipe, the assumption might be made that the latter contributes nothing to damping near the core, so that the velocity of near parts of the disturbed stream would have essential significance. Choosing for this velocity the maximum v_m of the rectilinear profile, we obtain fig. 9 from the measurements. Seven of the nine cores available for comparison appear nearly to establish within the range of size a limiting value of $\frac{2a_2 v_m}{\nu} = 90$, approximately, for the production of recti-

linear flow in the stream. The remaining two, of larger size, point to a marked " knee " in the curve in the region of $\frac{a_2}{a_1} = 0{\cdot}15$, where the pipe would appear to take charge. That the measurements tend towards zero Reynolds numbers as a_2 is decreased is probably explained by the disturbances remaining of constant magnitude. It would be of interest to experiment with extremely small inlet disturbances, which could not be secured with the present apparatus, when the method might be

Fig. 9.



Limiting critical Reynolds numbers for parallel streamline flow.
(v_m=max. velocity in annulus.)

expected to give definite information on the damping external to a long cylinder of circular section. But it may be inferred from the present investigation that the damping due to an internal cylindrical wall is much less than that due to an external one.

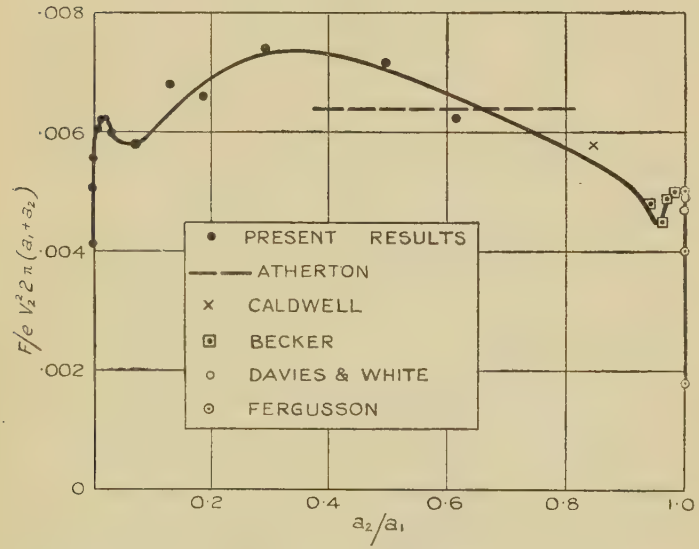
In regard to the second critical speeds, the simplest assumption to make is that " hydraulic mean depth " may prove a sufficient criterion. The broken-line curve of fig. 8 gives the variation with diameter-ratio of the quantity $\frac{v_{2'}a_1}{(a_1-a_2)}$, where $v_{2'}$ is the second critical speed of the pipe without a core in place. A rough

approximation is shown, the maximum difference being about 19 per cent.

It would be of greater interest to examine the records of second critical speeds in the light of the intensity of friction along the pipe-wall when they occur. This cannot be obtained, but the measurements give the mean intensity of friction along the whole channel at the critical speeds. It is convenient to introduce a resistance coefficient given by

$$k_D = \frac{F_2}{\rho v_2^2 \cdot 2\pi(a_1+a_2)}, \quad \ldots \ldots \quad (27)$$

Fig. 10.



Maximum resistance coefficient for persistence of turbulence through annuli.

where F_2 denotes the friction per unit length of the whole channel at v_2. Values of k_D are given in the last column of Table II. In fig. 10, where they are plotted, are included values derived from the experiments of Becker, Atherton, and Calwell, so far as these are obtainable from the published accounts, and also values of the same quantity observed with very narrow rectangular pipes by Davis and White * and by Fergusson † as applying

* Proc. Roy. Soc. A, vol. cxix. (1928).

† See Buckingham, 'Engineering,' vol. cxv. (1923).

approximately to the limiting case $\frac{a_2}{a_1} = 1$. All the observations appear to be fairly consistent. General turbulence sets in through all annuli, it appears, when the mean resistance coefficient lies somewhere within the range 0·0045–0·0074. This result may be compared with the value 0·0045 for curved pipes found by White* to be more or less independent of curvature.

The Strip as Core.—The strip was of bright spring steel, of width $w = 0{\cdot}373$ cm. and 0·004 cm. in thickness. The resistance of the channel formed between the pipe and strip was observed in the same way as for the annuli. Parallel streamline flow was established as far as a Reynolds number = 1050, specifying length, as before, by the diameter of the pipe. This value is more than twice as great as with a cylindrical core of equal contour and about the same as with one of equal width. Between the first critical Reynolds number and the second, which was found to be 2520, the index s had, within small experimental errors, the constant value 0·87.

The circumference of the pipe being about five and a half times the total contour of the strip, the effect on friction of the difference between its actual circular section and an assumed elliptic section confocal with the strip and having an equal area may be neglected, and comparison may then be made between the mean intensities of friction along circular and thin flat cores at equal mean velocities. From Section I. it is calculated that the ratio of the perimeter of the strip to that of the circular core whose resistance in streamline motion would equal that of the strip at the same mean velocity is 0·83, approximately. The experimental results below the first critical speed $\left(\text{mean } \frac{F}{\mu v w} = 44{\cdot}6\right)$ gave for this ratio the value 0·75. The difference may be ascribed chiefly to eccentricity, which was difficult to avoid along so great a length of steel strip, but a part may have been due to a rounding of its edges.

In turbulent flow evaluation of the above ratio is uncertain owing to an unknown difference in the effects of the strip and of a circular core on the resistance of the pipe. The mean resistance coefficient to correspond

* Proc. Roy. Soc. A, vol. cxxiii. (1929).

with column 5 of Table I. was 0·00670, increasing at higher speeds to 0·00687. The circular core of diameter-ratio 0·183 was specially tested to compare with the strip as having nearly the same perimeter, and this gave a corresponding resistance coefficient at higher speeds of 0·00751. Assuming the pipe-wall resistance to be unchanged by the cores, the ratio of the resistance of the flat core to that of the circular one was 0·65.

Thus it appears from these measurements that the mean intensity of skin friction along a thin flat strip is considerably less in streamline flow than that along the exterior of a cylindrical surface, while the effect probably also exists in turbulent flow. A practical, as well as theoretical, interest suggests the desirability of investigating this result by other means in less restricted flows.

The Visual Experiments.—These were undertaken to throw light on the cause of the early critical speeds. A glass tube, 4·57 cm. internal diameter and 35 diameters in length, was arranged with its axis vertical and its ends projecting into small tanks. The upper tank, open to the atmosphere, was fed with water over a long weir in one side, while a similar wier in the opposite side maintained the water at constant head. To regulate the flow the lower tank was fitted with an adjustable needle-valve, specially designed for easy flow, and drawing from a cylindrical rose coaxial with the glass tube. A round core of 0·475 cm. diameter was fitted approximately centrally within the tube, its ends projecting well beyond the tube, the upper end above the free surface of the water. Colouring fluid of approximately the same density as water was introduced thinly at the inlet of the annular channel.

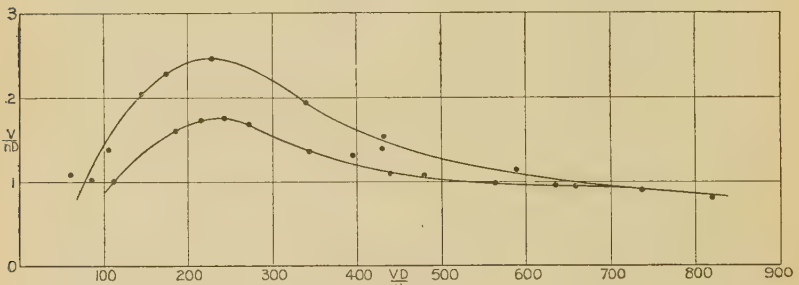
At low velocities the colouring fluid, at whatever radius introduced, continued as a dark line parallel to the axis. The photograph (*a*) of fig. 11 (Pl. XIV.) was taken just beyond the limit for rectilinear flow, and shows below the core a wave of small amplitude along the tube. A slight increase of velocity produced an easily seen slow sinuous motion, approximately a circumferential swaying, which further developed through a wide range of increase of flow through the annulus before breaking up into incoherent eddying. An advanced stage is shown at (*b*), fig. 11 (Pl. XIV.); photograph (*c*) was taken at a mean velocity

such that the thread of colouring fluid tended to break up ; (*d*) shows a still later stage.

The waving thread in (*b*) remained approximately at constant radius so far as could be seen. The axis of the oscillation of the thread wandered mostly in a circumferential direction, but sometimes radially ; it never circulated round the core. To verify absence of circulation drops of oil of various sizes were introduced in turn into the stream, and were maintained at constant level by adjusting the flow to their size. Oil-drops thus kept under observation for long periods through a wide range of speed showed no circulation to exist.

By fixing attention on a particular position along the tube, and observing the motion of the thread of colouring

Fig. 12.



Frequency of sinuous motion.

matter there, it was found possible to form consistent estimates of the mean frequency of the oscillation. The frequency was variable with time, as will be expected from the photographs, but average values estimated by counting through intervals of three minutes were found to have one or other of two values at each Reynolds number. The observations are given in fig. 12 in the form $\frac{v}{n\mathrm{D}}$ plotted against $\frac{v\mathrm{D}}{\nu}$, v being the mean velocity, n the number of complete oscillations per sec., and D the diameter of the tube. The sinuous motion only occasionally existed towards the upper extremity of the range in the figure, so that its persistence undeveloped through the present short tube extended, roughly, to the limiting Reynolds number for rectilinear flow observed with a long annulus of the same proportions.

It seems reasonable to assume that the phenomena observed with the short tube apply qualitatively to the flow through the inlet half of the stilling length of a long annulus of the wide kind. We infer, then, that shortly after entering such an annulus the bulk of fluid, unless the Reynolds number is less than 50 in the case of a diameter ratio of about 0·1, receives a swaying oscillation, a principal cause of which may well be variation of eccentricity, which is commonly present in experimental annuli, particularly near the inlet. This secondary motion has, presumably, to be damped out together with inlet disturbances for measurements of resistance past the stilling length to accord with the theoretical friction of rectilinear flow. It may be that the early failure of the core to produce damping is intimately associated with fluctuations of transverse components of velocity in its vicinity, when causes of disturbance exist.

Summary.

Steady viscous flow through long eccentric annuli is examined theoretically in some detail. On the restricted side of the channel the maximum velocity is shown to be notably less than on the opposite side. A small variable eccentricity being difficult to avoid in experimental annuli, lateral disturbances of the stream may commonly be expected. These are investigated visually in the case of a comparatively short annulus. They are suggested to be associated with an early failure to produce rectilinear flow in long annuli.

Nevertheless, in disagreement with earlier investigations, it is shown that up to certain Reynolds numbers—with initial disturbances—the familiar rectilinear flow for concentric annuli is established with a reasonable stilling length of channel. There are two critical speeds, the second being greater than for the pipe alone, and their variations are studied. With wide annuli the first is suggested to be due essentially to the core, the second mainly to the pipe-wall. A core only begins to delay to a marked extent the development of turbulence in flow through a pipe when its diameter exceeds half that of the pipe. In turbulent motion resistance increases much less quickly as the width of the annulus is reduced than in rectilinear flow.

A mathematical solution is obtained for steady flow through a channel bounded by confocal elliptic walls. When the sectional area of the core exceeds one-third that of the pipe, the proportionate increase for constant flux of the traction along the outer boundary due to the introduction of the core approximates, over a considerable range of mean width of channel, to that in the annular case. Particular attention is paid to the limiting case of a flat strip as internal boundary, and experiments are carried out for comparison. Comparing a thin flat strip in a confocal pipe with a circular core of equal resistance in a circular pipe, the pipe-walls enclosing equal areas, the ratio of the contour of the strip to that of the circular core, for equal discharge, is greater than unity, but less than $4/\pi$ in streamline flow. Experiments show this difference to be important also in turbulent flow. Introducing a particular strip into a long pipe reduced by one-half the range for streamline flow, but this range remained about twice as great as that with a circular core of equal perimeter.

The authors desire to thank Professors L. Bairstow, C.B.E., F.R.S., and C. H. Lees, F.R.S., for valuable advice; Mr. S. L. Green, M.Sc., for reading the mathematical part of the work; and the University of London for certain financial provisions from the Dixon and other funds.

LIV. *On Deposits of Metallic Mercury by High-frequency Discharge.* *By* D. Banerji, *M.Sc., Lecturer in Physics, University College of Science, Calcutta, and* Radharaman Ganguli, *M.A., M.Sc., Lecturer, Serampore College, Bengal.*

[Plate XV.]

Introduction.

In course of some experiments on high-frequency discharge in gases, it was noticed that certain greyish deposits in the shape of rings were formed alongside the external sleeve electrodes which fed the high-frequency current to the discharge-tube. As we have not been able

*. Communicated by Prof. S. K. Mitra, D.Sc.

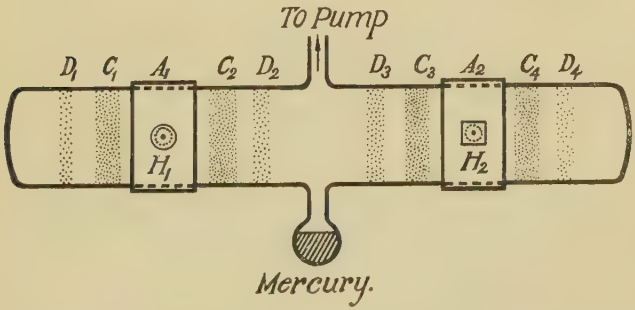
to find on record any mention of such deposits by investigators in this line, we were led to make a preliminary study of the nature and origin of such deposits. The results of our study with different shapes and sizes of the discharge-tube and with different dispositions of the electrodes, are set forth in this paper.

It has been found that the deposit is mainly of mercury, which obtains access to the discharge-tube from the McLeod gauge. The failure to observe the deposit by other workers is probably due to the fact that care is generally taken to prevent mercury from reaching the main discharge by interposition of liquid-air traps. An examination of the distribution of space potential inside the tube gives an indication of the mode of formation of the deposits.

Experiment and Observation.

The deposit can be obtained quickly and easily by passing high-frequency discharge through an evacuated

Fig. 1.



glass tube containing distilled mercury in a side-bulb. In an actual experiment a glass tube, 20 cm. long and 2 to 3 cm. wide, with a side-bulb was used as shown in fig. 1. The sleeve electrodes used were each 2·5 cm. wide, separated from each other by a distance varying from 6 to 16 cm. The frequency of the exciting potential was about 4 million This was obtained in the usual manner as used in our previous experiments *.

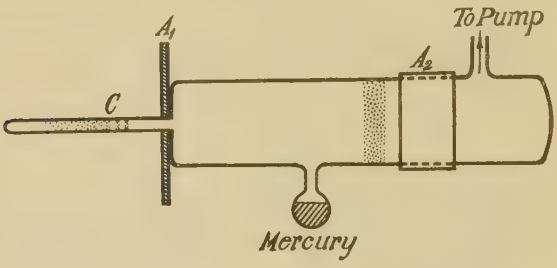
To start with, the mercury in the bulb was slightly heated to produce sufficient mercury vapour. Four

* Phil. Mag. xi. pp. 410–22 (Feb. 1931).

ring-like deposits, C_1, C_2, C_3, and C_4 (fig. 1), appeared close to the four edges of the electrodes, A_1, A_2, after a quarter of an hour. These transmitted the green light of mercury and became opaque after a run of about two hours. During this interval four other fainter rings, D_1, D_2, D_3, and D_4 (fig. 1) of deposit were seen to grow at a distance from the main rings.

In actual experiments the first set of inner rings were formed at a distance of 5 mm. from the edges of the electrodes, and the outer rings at a distance of $2\frac{1}{2}$ cm. The deposits always seem to avoid the immediate neighbourhood of the electrodes, and are never formed beneath or on the edges of the electrodes. If a hole is cut open in the electrodes, the deposit is formed round the centre of the holes H_1 and H_2, but

Fig. 2.



it never reaches the boundary. All the above features, namely, a pair of heavy rings and a pair of faint rings as well as the nature of the deposit inside the hole cut in the electrode, can be easily seen in figs. 1 & 2 (Pl. XV.) particularly in the latter, which is a close up view of the deposit near one of the two electrodes.

The deposit can also be obtained more quickly by the arrangement shown in fig. 2. This is a glass tube with an extension towards one side by a narrow tube (diameter 5 mm.). One of the external electrodes is a flat circular piece with a central hole, A_1 (fig. 2), fitting into the narrow tube. The other one, A_2, is of the usual sleeve form. At a particular pressure (·8 mm. of mercury in the actual experiment) the discharge penetrates into the narrow tube like a pointed tongue, and the deposit grows in the form of short rings C (fig. 2) close to the tip of the tongue in a few seconds.

Attempts were also made to obtain the deposit by fitting electrodes inside the glass tube close against the wa ls. It was found that in such a tube the deposit took a comparatively long time to form.

Experiments were also made by passing the discharge through iodine vapour. In this case also the iodine was deposited on the walls of the tube, but the deposit was uniform, and no ring-like form, as in the case of mercury, could be noticed.

The deposits may also be formed by the mercury vapour finding its way to the discharge-tube from the McLeod gauge even when there is no mercury in the side-bulb. In this case, however, the deposits take a longer time to form. With mercury in the side-bulb, slightly heated to start with, one obtains an observable deposit in about ten minutes time. In order to get the same density of deposit without the mercury in the side-bulb, but with the McLeod gauge only, it takes about an hour.

That the deposit is due to the presence of the mercury vapour is conclusively proved by cutting off the McLeod gauge. No deposit can be obtained in this case even after a run of three hours.

The deposit when examined by breaking open the tube is found to stick firmly on the inner walls of the glass tube, and to present gorgeous interference colours of thin films. These films can partially be removed by rubbing, or by prolonged heating over a gas flame. They also dissolve in boiling nitric acid. A faint trace of the deposit, however, remains and refuses to dissolve after prolonged heating, probably owing to the fact that the mercury atoms penetrate into the walls of the glass tube.

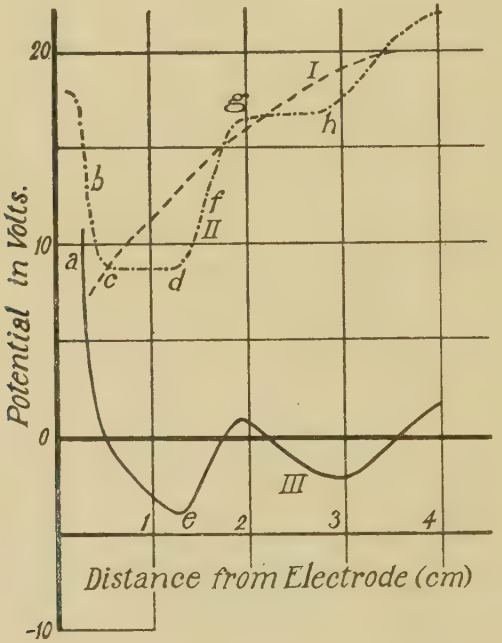
Discussion.

It is natural to suppose that the deposit is due to the presence of charged mercury particles which move up to the glass walls under the action of the electric field produced in the tube by the high-frequency discharge. One might therefore enquire into the nature of the potential distribution inside the tube which would cause the charged particles to move in such a manner as to arrive at those particular positions on the glass walls where the deposits are formed.

With this end in view the distribution of space potential was measured both along the axis and a so along the length of the wall of the glass tube by the Langmuir collector method, with the help of a third electrode as described in our previous paper *.

In fig. 3. Curve I. shows the variation of the space potential along the axis of the tube and Curve II. that along the wall of the tube, the distances being measured

Fig. 3.



from the left electrode. It will be observed that while the potential along the axis gradually rises to a maximum near the centre, that along the wall starts from a maximum near the electrode, drops to a minimum value at a distance of about 1 cm. from the electrode, and then rises by a number of steps as shown in Curve II. Curve III. shows the difference between the potential at a point on the wall and the corresponding point on the axis. This curve therefore shows the rise and fall of electric force acting in the radial direction as one proceeds along the axis

* *Loc. cit.*

from the left electrode. If we consider a positively charged particle situated near to the point *a* on the axis, the particle will experience a radial force urging it towards the wall. When it reaches the vicinity of the wall, the electric force along the wall, which is a maximum at *b*, will tend to drag the positive particle along the wall. The particle will thus hit the wall in a region between *c* and *d*, where the electric force along the wall is a minimum. If the particle is negatively charged then it will experience a maximum radial force when it is on the point *e* (Curve III.). It will move up towards the wall, but when nearing the wall it will experience an electric force along the wall which has a maximum value at *f* (Curve II.). The particle will thus hit the wall between the regions *g* and *h*. In the photograph (Pl. XV. fig. 1) it may be noticed that the position of the deposit corresponds tolerably with the position of the regions *cd* and *gh* in Curve II. The heavy deposit near the electrode is due to positively charged mercury ions, *i. e.*, mercury atoms from each of which one electron has been knocked out, while the deposit further from the electrode is due to negative mercury ions or mercury atoms to each of which an electron has got itself attached.

Summary and Conclusion.

It has been found that under certain conditions sets of coloured rings are formed on the inside wall of a high-frequency discharge-tube containing mercury vapour. These rings, which are due to deposit of mercury, appear close to the electrodes fitted externally to the tube. Its mode of formation is explained by the distribution of space potential along the axis and along the wall of the tube. Photographs depicting the shape and position of the rings are given.

The phenomenon was noticed by Prof. S. K. Mitra some years ago, when studying the excitation of a Cooper-Hewitt mercury lamp by high-frequency discharge. The observation was communicated to us by him, and the present investigation was undertaken at his suggestion. We take this opportunity of thanking him for his helpful interest and advice during the course of the work.

Wireless Laboratory,
University College of Science,
92 Upper Circular Road, Calcutta, India.
Sept. 26, 1932.

LV. *On the Theory of Oscillatory Condenser Discharges.* *By* JOHN THOMSON, *M.A., Ph.D., D.Sc., Lecturer in Natural Philosophy at the University of Glasgow* *.

[Plate XVI.]

Introduction.

THE purpose of the present communication is the further elucidation of condenser discharges. Many attempts have been made to explain the sequence of events which follow the break-down of the insulation of a small air-gap in parallel with a condenser which is charged to a high potential, and the broad features of the phenomena are well understood. In the first part of this paper experiments are described which exhibit this sequence of events very clearly, and which suggest a theory of the condenser oscillations. This theory is explained in detail, and it is shown that it is consistent with all the facts known to the writer. In the second part of the paper a series of similar experiments dealing with the oscillating arc is described. It is shown that the explanation of the behaviour of the arc is analogous to the explanation of the mechanism of the oscillating "spark," and, indeed, that a more exact treatment of the latter may be obtained from a consideration of the former. A theory of the different forms of oscillating arc is then sketched, and it is shown that this theory is in agreement with experiment. The fundamental conception of the theory is that the potential across a gaseous discharge the current through which is varying may always be represented as a function of the time consisting of the sum of two exponential terms $Le^{-ht}+Me^{ht}$, so long as the discharge is self-sustained. It is also shown that this theory satisfactorily accounts for the oscillating condenser "spark."

An important conclusion which may be drawn from the experiments is that there is no fundamental distinction between the later stages in a *high* potential discharge and the normal type of "singing" arc, any difference being of degree rather than kind. It is also suggested that the use of the term "inductance of an arc" is very

* Communicated by the Author.

liable to mislead. There is no experimental evidence of an inductance being associated with a discharge at the frequencies investigated by the writer.

Part I.—*The Mechanism of the "Spark."*

The writer has already described certain experiments on condenser discharges [1], and has suggested that the initial stage in the break-down of any gaseous dielectric may be regarded as a condenser "spark." It has been shown (both in the earlier communication and by the present experiments) that when the electrostatic charge on a condenser is liberated through an air-gap, the initial "glow" very rapidly degenerates into a typical "arc," and that it is this arc-like phase which gives rise to the spark spectrum of the metal of the electrodes.

The attention of the writer was recently called to a paper of Professor Milner's entitled "The Current-Potential Curves of the Oscillating Spark and the Mechanism of Spark Conduction" [2], which contains remarkable evidence in favour of the above contention. It was thought advisable, therefore, to repeat Professor Milner's experiments, utilizing the more refined technique of the present day, and, as will be seen later, the repetition has certainly proved to be worth while.

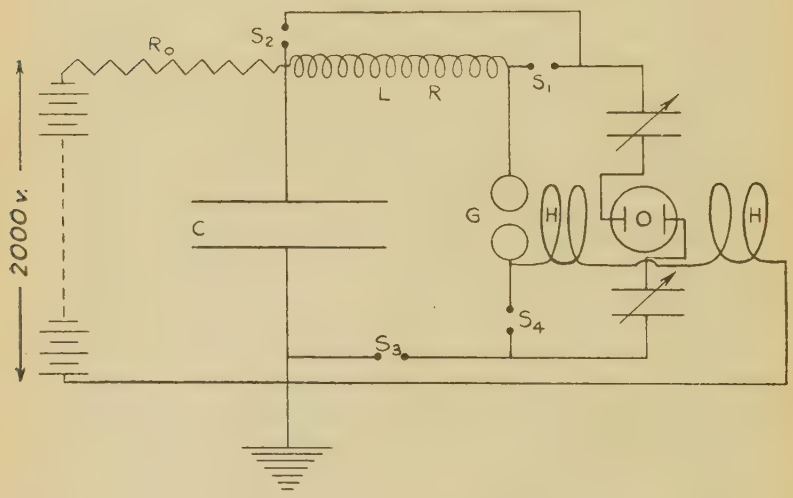
It is well known that the fundamental difficulty in the way of an investigation of the mechanism of a condenser discharge is the rapidity with which the various stages in the phenomenon follow one another. For example, if a condenser of capacity $1\mu F$ is discharged through an inductance of $1\mu H$ and an air-gap, the oscillations have become of negligible amplitude in a time less than 10^{-5} sec. With modern "time-base" technique employed in conjunction with a cathode-ray oscillograph, it is certainly possible to examine these oscillations, but, even with all modern improvements the details of the oscillograms obtained at such high frequencies are difficult to interpret. Professor Milner's experiments obviated the necessity for a time base by simultaneously deflecting the cathode ray beam by means of an electric field proportional to the voltage across the condenser (or spark-gap) and a magnetic field proportional to the current flowing through the spark-gap, the two deflexions being at right angles. This method, which has been employed many times since

in other investigations, exhibits the current-potential characteristic of the system under observation. It is surprising that it has not (to the writer's knowledge) been more freely used for the examination of discharge processes.

Experimental Arrangements.—The present writer's circuit is shown diagrammatically in fig. 1.

A high potential was produced across the condenser C (1μF) by means of a battery of dry cells giving approximately 2000 volts. The positive pole of the battery was earthed, and the negative pole was connected to C

Fig. 1.



through a large protective resistance R_0. The negative pole of C was connected through an inductance L (0·08 henries) to the spark-gap G (zinc hemispheres, 2 cm. diam.). The other side of the gap was connected to the earthed terminal of the condenser through a Helmholtz coil H, which was arranged to produce a uniform magnetic field at the appropriate region of the oscillograph O. The latter was a Cossor instrument of the standard type in which all four plates are insulated from the gun. Two opposite plates were connected to two small variable air condensers, the other leads from which were taken to non-capacitative switches S_1, S_2, S_3, S_4. Thus the two oscillograph plates formed the middle condenser

of a system of three in series ; and this system could be connected by switches S_1, S_3 in parallel with C, or, by switches S_2, S_4 in parallel with G. The other two deflecting plates of the oscillograph were connected together and earthed.

When using a modern oscillograph in the manner described above, the deflexion being due entirely to the induced electrostatic charge on the plates, it is important (if the frequency of the applied voltage is less than 10^6 per sec.) to verify that the ionic current to the plates is negligible. If this condition is not fulfilled, the readings of the instrument are incapable of exact interpretation. Fortunately, in the present experiments it was possible to test the action of the oscillograph directly by means of a known volt-ampere characteristic, and it was found (after numerous other attempts) that the above circuit was reliable. With the arrangement described it is probable that almost all the ionic current flows to the pair of deflector plates, which are connected together and earthed. An almost similar effect can be produced (if low-frequency oscillations are to be examined) by connecting three high resistances in parallel with the condensers, and this latter circuit has the advantage that a steady potential across the condensers produces a proportional steady deflexion of the fluorescent spot on the screen. Both arrangements were used in the experiments to exhibit different effects.

The Helmholtz coil H was placed so that a current through it produced a deflexion of the beam in a direction perpendicular to the voltage deflexions ; the usual precautions were taken to eliminate stray magnetic, electrostatic, and electromagnetic fields.

Results.—With the circuit of fig. 1 the gap-length was gradually reduced until regular " sparking " occurred. The switches S_2 and S_3 were then closed so that the voltage deflexions shown by the oscillograph were proportional to the potential difference across the condenser C, and the current deflexions were proportional to the current flowing in the condenser to spark-gap circuit. As was to be expected, the resulting oscillogram was (approximately) an elliptical spiral in agreement with the Kelvin theory, and in agreement with Professor Milner's result. Photographs of such patterns are shown in figs. 1, 2, 3 (Pl. XVI.). These were taken with a camera of focal length 10 cm. and aperture

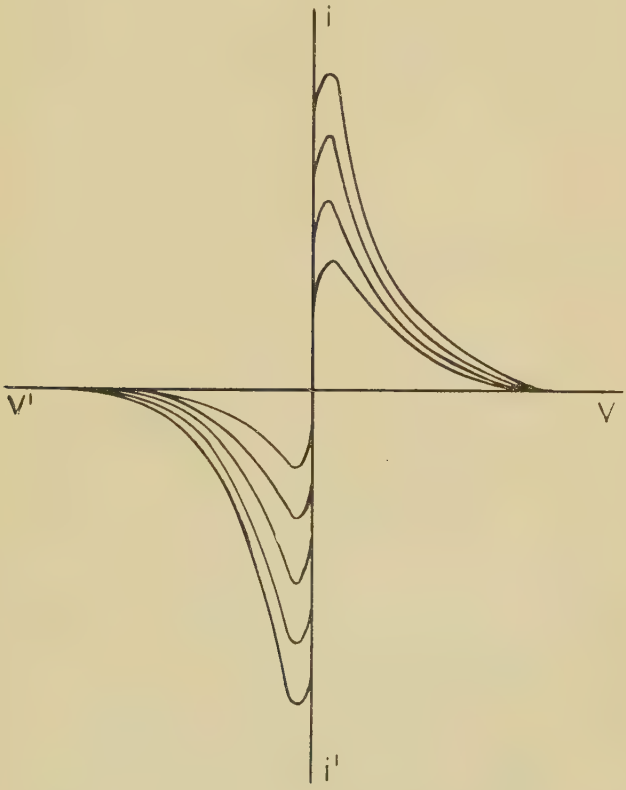
5 cm., using Golden Izo-Zenith Plates, but the pattern was so faint and was traced so rapidly that many "sparks" had to be passed to obtain a useful negative. However, the definition is reasonably good. Fig. 1 (Pl. XVI.) is the standard pattern. The voltage axis is represented by the line VV′, so that the major axis of the ellipse is inclined at an angle of about 40° to the V-axis. This indicates a considerable damping in the oscillatory circuit. In fig. 2 (Pl. XVI.) a *very* large resistance was placed in parallel with the oscillograph plates. The net effect of this was to superimpose upon the oscillating voltage an aperiodic decaying voltage, so that the ellipses move progressively towards the origin as the oscillations die out. This method was adopted to eliminate the effects due to the glare of the central spot. The voltage axis is again indicated by VV′. The circuit for fig. 2 (Pl. XVI.) contained less inductance than that for fig. 1 (Pl. XVI.). Hence the angle between VV′ and the major axis is greater in fig. 2 (Pl. XVI.), the ratio R/L being increased.

A careful examination of figs. 1 and 2 (Pl. XVI.) shows quite definitely that the patterns are not continuous. On the voltage axis a sharp break takes place : partly due to the oscillograph screen being less actively fluorescent along the axis, but partly due to a real effect of the circuit. To test this various combinations of L, C, R were used with various spark-gaps, and fig. 3 is a pattern obtained with one combination. Here the ellipses are distorted, due to a rectifying action in the gap, but, quite clearly, the pattern is discontinuous along VV′. This discontinuity is of considerable importance in the theory of condenser oscillations. Its explanation will be given later.

Using the circuit with which fig. 1 (Pl. XVI.) was obtained, switches S_2 and S_3 were now opened, and S_1, S_4 were closed. The voltage deflexions shown by the oscillograph were now proportional to the potential difference across the spark-gap, the current deflexions remaining as before. A typical result is shown in fig. 4 (Pl. XVI.), but the reproduction of the original negative fails to show some of its more interesting features. This figure is the volt-ampere characteristic of the oscillating spark, and it is in complete agreement with Professor Milner's characteristic. It is clear that the large current flows through

the gap at a small voltage, and that there is no obvious resemblance between the condenser characteristic and the gap characteristic. The original screen pattern showed much more detail than the photograph, and fig. 2 indicates these details.

Fig. 2.



Theory of Condenser Oscillations through an Inductance, Resistance, and Air-Gap (Part I.).

The Kelvin theory of condenser oscillations is derived from the solution of a differential equation of the type

$$\frac{d^2V}{dt^2}+a\frac{dV}{dt}+bV=0,$$

i. e., the condenser is discharged through a constant

resistance and inductance. It is well known that the resistance of a spark-gap is a very variable quantity, and many attempts have been made to express it in terms of some other variable such as the current flowing in the gap, or, as in the case of Toepler's formula, the total quantity of electricity which has previously been discharged.

The writer has attacked the problem in the following different manner. Suppose that the potential difference across the gap at time t is $V_s(t)$. Then the differential equation for a circuit such as that used in the experiments is

$$V = Ri + L\frac{di}{dt} + V_s(t),$$

where i is the current, V the condenser voltage, R and L the resistance and inductance.

But $$i = -C_1\frac{dV}{dt}$$

where C_1 is the condenser capacity.

Therefore

$$\frac{d^2V}{dt^2} + \frac{R}{L}\frac{dV}{dt} + \frac{1}{LC_1}V = \frac{1}{LC_1}V_s. \quad . \quad . \quad . \quad (1)$$

This equation may be solved explicitly if V_s is known in terms of t.

From Professor Milner's results and an examination of oscillograms taken by Roschansky [3] it is very probable that a good approximation to V_s is

$$V_s = A + Be^{-kt},$$

where k is large. It will be shown later that this equation may be modified by the inclusion of another exponential term Ce^{kt}, or, by changing A to Ae^{kt}, but a first approximation is given by the simple equation. Suppose, therefore, that V_s varies between V_G and V_A, two constant voltages, then

$$V_s = V_A + (V_G - V_A)e^{-kt} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

This holds for one half of the condenser oscillation. In the other half-oscillation

$$V_s = -\{V_A + (V_G - V_A)e^{-kt}\},$$

but this is a deduction from the theory rather than an empirical statement. The solution of equation (1) then becomes

$$V=V_0e^{-\frac{R}{2L}t}\cos(pt-\delta)+V_A+\frac{V_G-V_A}{n^2+1-\frac{nR}{pL}}e^{-npt}, \quad . \quad (3)$$

where $$p=\sqrt{\left\{\frac{1}{LC_1}-\frac{R^2}{4L^2}\right\}}$$

and $$k=np.$$

If n is large,

$$\frac{V_G-V_A}{n^2+1-\frac{nR}{pL}}$$

is small compared with V_0,

and

$$V \doteqdot V_0e^{-\frac{R}{2L}t}\cos(pt-\delta),$$

which is the simple solution.

This solution agrees well with the observed characteristics, but its chief interest lies in the investigation of the phenomena occurring when the discharge stops temporarily at $pt\doteqdot\pi$. The writer's suggestion is that, when the discharge goes out, the circuit behaves as if it consisted of two condensers, one large and the other small, coupled by the inductance. It therefore oscillates with a very high frequency until the redistribution of electricity makes it possible for the discharge to recommence.

To simplify the equations, it may be assumed without loss of generality that R/2L is negligible in comparison with p. Then (3) becomes (if V_0 is large)

$$V=V_0\cos pt+V_A+\frac{V_G-V_A}{n^2+1}e^{-npt}.$$

At $pt\doteqdot\pi$, $i=0$ and $V=-V_0+V_A$.

Also $$L\frac{di}{dt}=-V_0.$$

It is at this point in time that the discharge goes out and the gap becomes non-conducting. Fig. 3 then represents the new circuit where C_1 is the capacity of the condenser and C_2 is the capacity of the gap.

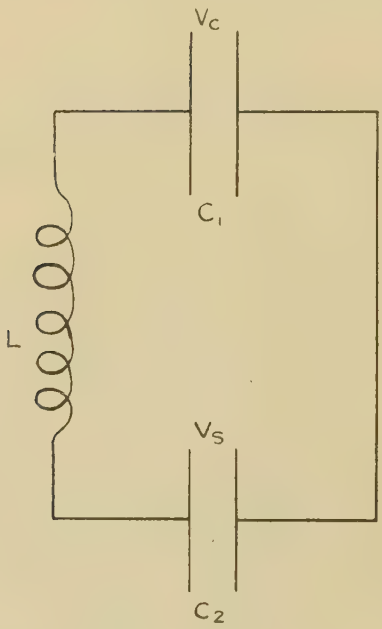
The equations are, using $t'=t-\frac{\pi}{p}$.

$$V_c=L\frac{di}{dt'}+V_s$$

$$i=-C_1\frac{dV_c}{dt'}=C_2\frac{dV_s}{dt'}.$$

Let $$i=A\sin p't'.$$

Fig. 3.



Therefore, since

$$\frac{di}{dt'}=-V_0, \quad \text{when } t'=0, \quad A=-\frac{V_0}{Lp'}$$

and

$$V_s=V_A-V_0(1-\cos p't')$$

while $$V_c=V_A-V_0+\frac{V_0p^2}{p'^2}(1-\cos p't').$$

Hence, in a time which is negligible compared with $\frac{2\pi}{p}$ (the period of the circuit when the discharge is in

action) V_s becomes $-V_G$, a potential at which the discharge may recommence. This occurs when

$$\cos p't' = \frac{V_0-(V_G+V_A)}{V_0},$$

and at this time,

$$V_c \doteqdot -(V_0-V_A').$$

As a matter of fact, it is unlikely that the discharge recommences when V_s becomes $-V_G$ in the first oscillation. According to the condition of the gap, one, two or more high-frequency oscillations may take place before the discharge begins. It is noteworthy that V_s oscillates between the wide limits, V_A+V_0 and $-(V_0-V_A)$, while V_c only oscillates between

$$-(V_0-V_A)+p^2V_0/p'^2 \quad \text{and} \quad -(V_0-V_A)-p^2V_0/p'^2.$$

Since p/p' is very small, this oscillation is of negligible amplitude.

A Test of the Theory.—Attention has already been called to the discontinuities on the voltage axis in the oscillograms of figs. 1, 2, 3 (Pl. XVI.). The explanation of these discontinuities is now apparent. When the discharge ceases, the circuit oscillates with a much higher frequency, and the breaks in the low-frequency pattern represent these rapid oscillations. Similar discontinuities are faintly visible in several of the oscillograms reproduced in the paper by Roschansky, but the latter does not attempt to explain them.

To test this theory a modification of the oscillograph circuit was arranged. If a resistance is placed in series with the oscillograph across the part of the circuit in which the voltage fluctuations are to be observed, then the capacity of the oscillograph charges up relatively slowly through the resistance. It is possible, therefore, to choose the value of the latter so that the oscillograph will respond to oscillations of low frequency, but will pay no attention to high-frequency effects. With this modified circuit it was found that the spiral patterns showed little or no interruption on the voltage axis. The high-frequency nature of the discontinuities is thus demonstrated.

Conclusions.—From the experiments just described the following conclusions may be drawn :—

1. It is quite certain that a capacity discharge is, for

the greater part of its life, a low-potential, high-current "arc."

2. The discharge ceases to be self-sustained when the potential across the gap becomes zero, and therefore the discharge ceases twice in every oscillation of the condenser circuit.

3. This makes it necessary to assume that oscillations of two frequencies occur in the condenser circuit. The normal oscillation of pulsatance $p \doteqdot \frac{1}{\sqrt{(LC_1)}}$ (where C_1 is the capacity of the condenser and L the series inductance) is characteristic of the circuit when the discharge is passing. When the discharge goes out, a high-frequency oscillation of pulsatance $p' \doteqdot \frac{1}{\sqrt{(LC_2)}}$ (where C_2 is the capacity of the spark-gap) takes place, redistributing the energy in the circuit so that another discharge may begin.

4. In this connexion it may be suggested that the comparatively low potential V_G at which the discharge subsequently begins is due to the fact that the electric field in the gap is a high-frequency one, it being well known that a high-frequency glow may be initiated at a very much lower potential than a direct-current one.

5. It is probable that the potential across the spark-gap while the discharge is running may be given fairly accurately by an expression

$$V_s = V_A + (V_G - V_A)e^{-npt},$$

where V_A is the gap-potential at which the discharge goes out and V_G is the gap-potential at which the discharge recommences. If the suggestion of the last paragraph is not accepted, V_G may be regarded as a dynamic sparking potential, very much lower than the static sparking potential. The factor of t in the exponential (np) is, of course, positive.

6. If V_s is given by an expression of the above type, the experiments indicate that n is very large—of the order 100. This is necessary in order that the condenser volt-ampere characteristic should remain approximately a spiral.

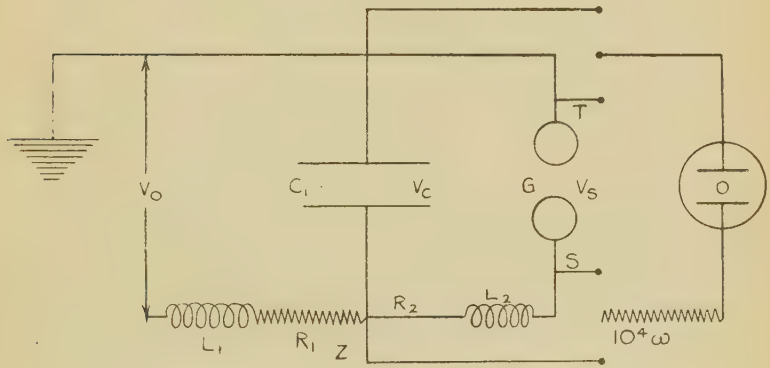
Part II.—*The Mechanism of the Oscillating Arc.*

The previous investigation suggested to the writer that the phenomena associated with the later stages in condenser "sparks" might be realized by means of a circuit

giving rise to a "singing arc." The argument was, briefly, that if the current through the discharge in the circuit of fig. 1 was increased by decreasing R_0 and having the gap G initially closed, then, on opening the gap slightly, it might be possible to examine in more detail the low-potential, high-current part of the oscillogram of fig. 3 (Pl. XVI.). With this arrangement it would be possible (it was argued) to connect the oscillograph plates directly across the gap G, thus increasing enormously the voltage sensitivity. It was hoped that the investigation would test and possibly suggest improvements in the theory given in Part I.

Experimental Arrangements.—Essentially, the new circuit

Fig. 4.



(fig. 4) was similar to the old, but the values of the resistances and inductances were altogether different.

L_1 was about 0·08 henries, R_1 was variable between 50 and 250 ohms, R_2 represents only the unavoidable resistances of the leads and electrodes, while L_2 was small, about 30 microhenries. It was found that the potential across the electrodes sometimes rose to as much as 250 volts, so a protective resistance of 10,000 ohms was placed in series with the plates of the oscillograph 0 as shown. The capacity of the oscillograph plate system is not greater than $12\mu\mu$F., so that the plates acquired 99 per cent. of the value of a steady potential applied across G or C_1 in 6.10^{-7} sec. This time was found to be very small compared with the period of oscillation of the circuit, and so, in spite of the theoretical inaccuracy

of the system, the arrangement of a resistance in series with the oscillograph was retained. As before, part of the inductance L_2 was arranged in the form of a Helmholtz coil about the axis of O to produce deflexions of the cathode ray beam proportional to the current. This part of the circuit has been omitted in fig. 4 for the sake of simplicity.

The capacity C_1 was, usually, a 2 microfarad condenser. The gap G was between the two electrodes of a hand-feed arc lamp, and special precautions were taken to see that the presence of this steel frame near the oscillograph did not produce errors.

Results.—In the first experiment two carbon electrodes were used. It was found, however, that the resistance of the pieces of carbon between the leads S and T and the gap was sufficient to produce a measurable E.M.F. across the oscillograph. Therefore the positive carbon electrode was replaced by a solid copper cylinder. This allowed the other carbon to be made very short, so that the resistance between S, T, and the gap was reduced to a minimum.

Copper-Carbon Arc.—Fig. 5 (Pl. XVI.) exhibits the mode of variation of the current in the arc circuit (R_2L_2G) with the potential across C_1. The exposure in this case was about one-tenth of a second, so that many oscillations of the circuit are traced on the photograph. The rectifying action of the arc is the first feature of the oscillogram to catch the attention of the observer. Practically no current flows in the negative direction, although, since ii' represents the current axis, the potential across the condenser is negative during a considerable fraction of each oscillation. The form of the curve for positive current strongly suggests

$$V_c = A \cos pt + B$$

$$i = D \sin pt + E,$$

omitting (for simplicity) the probable decay factor.

Fig. 6 (Pl. XVI.) exhibits the mode of variation of the same current (circuit R_2L_2G) with the potential across the arc. Again the rectifying action of the gap is clearly demonstrated, and, as was to be expected, a typical arc characteristic is obtained. At the moment no more need be said of the features of this figure, but an additional illuminating experiment will be described.

To examine the nature of the current flowing in the L_1R_1 part of the circuit, an additional coil was inserted between R_1 and the point Z and arranged to take the place of the Helmholtz coil producing current deflexions in the oscillograph. In this way the mode of variation of the current in R_1L_1 was exhibited as a function of V_c, and it was found that this current was quite accurately given by

$$i_1 = \frac{V_0 - V_c}{R_1},$$

within the range studied. This result (which is only an approximation) will be found to be very useful in the discussion given later.

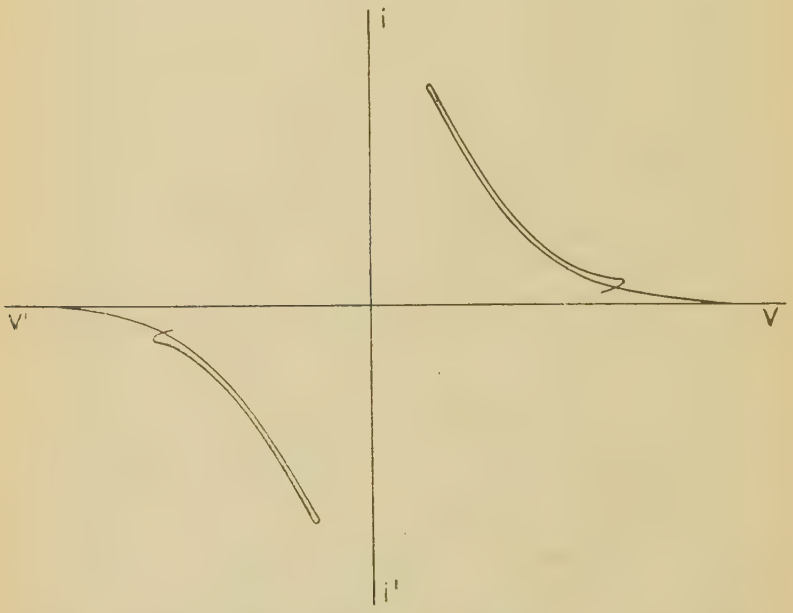
Copper-Copper Arc.—The experiments with the copper-carbon arc, while yielding interesting oscillograms, did not give direct evidence of the occurrence in the arc of any of the phenomena associated with the condenser "spark." It was, however, fairly obvious that this was due to the rectifying action of the combination. The negative carbon was therefore replaced by a copper cylinder also, and the experiments were repeated. In this case it was impossible to obtain a stable singing arc—when a periodic discharge *did* take place, it was quite different from the steady sustained oscillation of the other combination. The discharge (when periodic) was accompanied by a series of sharp reports, reminding one of recurring sparks. The oscillograms obtained from this arrangement showed little or no valve-action. The current was always less in one direction than in the other, but that could be traced to the fact that initially a large unidirectional current was required to start the oscillations. The condenser characteristic was roughly elliptical, while the important oscillogram, namely, the gap characteristic, is shown in fig. 7 (Pl. XVI.). The steady arc position is shown as a diffuse, but intense spot of light, the current oscillations as broad bands. The potential axis is the fine line across the photograph, showing that the potential oscillated while no current was flowing in the gap. From the nature of this oscillogram several important deductions can be made, viz. :—

(*a*) The characteristic is reminiscent of Professor Milner's pattern. The current axis just fails to touch the characteristic when the current is large; in fact,

the axis is almost an asymptote to the limiting curve of the negative part of the pattern. This indicates that the current through the gap may be very large for very small values of V_s, but is always zero when the potential is zero.

(*b*) The characteristic is definitely *not* in agreement with the simple theory sketched in Part I., for, as the current falls to zero, the potential certainly does not either decrease or remain approximately constant. The

Fig. 5.



oscillogram is more appropriate to the suggestion that, as the current decreases again after reaching its maximum value, the potential increases again after simultaneously reaching its minimum value.

Discussion of the Results.—Figs. 4, 6, 7 (Pl. XVI.) all represent what have been called "gap characteristics." Fig. 4 (Pl. XVI.) is on such a scale that it is impossible to examine in detail the high-current, low-potential part of the pattern. Figs. 6 and 7 (Pl. XVI.) both suggest, however, that they are obtained from the superposition of curves of the form shown in fig. 5. In other words, they

suggest that the potential across the gap, *during the time when the discharge is running*, first decreases and then increases as the current increases and then decreases. This is, of course, in agreement with the well-known formula for the arc characteristic

$$V_s = a + \frac{b}{i} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where i is the current; although the agreement so far is merely qualitative. The writer suggests, however, that in *all* capacity discharges, the sequence of events may be represented by characteristics of the form of fig. 5, unless in extreme cases where the inductance in the discharge circuit is very small and the current charging the condenser is comparatively large. He does *not* suggest that such characteristics are parts of hyperbolæ as indicated by equation (4).

Theory of Condenser Oscillations through an Inductance, Resistance, and Air-Gap (Part II.).

In modification of the theory given in Part I. the writer would now suggest that the potential across the discharge may *always* be represented by an expression of the form

$$V_s = \pm\{Le^{-kt} + Me^{kt}\}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where L, M, k are positive quantities depending upon the circuital constants and the initial conditions. This expression only holds, of course, for the time during which the discharge is self-sustained. In general a comparatively abrupt change in V_s takes place when the arc goes out.

V_s has a turning-point at the value of t given by T,

where $$kT = \tfrac{1}{2}\log_e\left(\frac{L}{M}\right). \quad . \quad . \quad . \quad . \quad . \quad (5')$$

At this turning-point by substitution from (5′)

$$V_s = 2\sqrt{LM}.$$

Hence, if the initial potential across the gap is known, say V_G; and the turning-value of the potential is known, say V_A, then L and M can be determined immediately from the equations

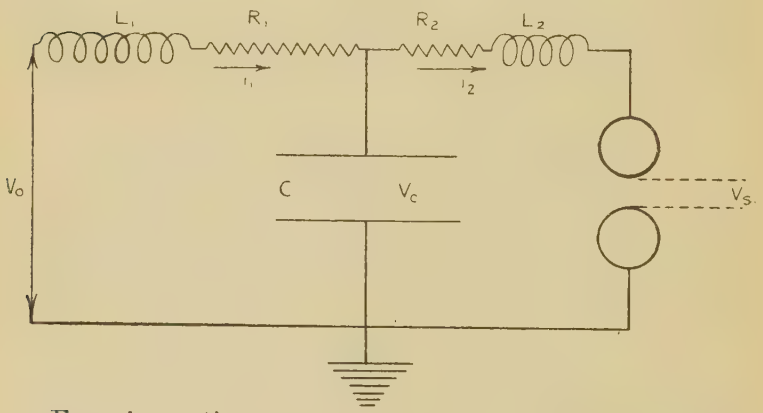
$$L + M = V_G,$$
$$LM = \tfrac{1}{4}V_A^2.$$

These equations also determine the value of the product kT, but, let it be noted, not the values of either k or T.

Hence, if V_s is represented by the sum of two exponentials as in equation (5), it is always possible to arrange that V_s shall have a turning-point at any predetermined value of t without prejudice to the predetermined values of the initial V_s and the turning-value of V_s.

Proceeding now to the examination of the circuital equations for the general case of condenser oscillations, fig. 6 is the circuital diagram.

Fig. 6.



From inspection

$$V_c = R_2 i_2 + L_2 \frac{di_2}{dt} + V_s.$$

and

$$i_1 - i_2 = C \frac{dV_c}{dt}.$$

To write down the equations expressing V_c in terms of i_1 and V_0 and to eliminate i_1 and i_2 leads to a differential equation of the 3rd order, the coefficient of $\frac{d^3V_c}{dt^3}$ being L_1L_2C.

Assuming $L_1=0$, leads to the equation

$$\frac{d^2V_c}{dt^2} + \left(\frac{1}{CR_1} + \frac{R_2}{L_2}\right)\frac{dV_c}{dt} + \frac{1}{CL_2}\left(1 + \frac{R_2}{R_1}\right)V_c = \frac{1}{CL_2}V_s + \frac{R_2}{CL_2R_1}V_0, \quad . \quad . \quad . \quad (6)$$

which is exactly the equation obtained if i_2 is put equal to $\frac{V_0-V_c}{R_1}$. This last assumption is, however, an experimental result, for it will be remembered that it was found that i_2 varied linearly with V_c in exactly this manner. Hence it is legitimate to assume that in the experiments described L_1 was too small to be of any importance. Writing equation (6) in another form and inserting the hypothetical value of V_s,

$$\frac{d^2V_c}{dt^2}+a\frac{dV_c}{dt}+bV_c = \frac{1}{CL_2}[Le^{-kt}+Me^{kt}]+\frac{R_2}{CL_2R_1}V_0. \quad . \quad . \quad . \quad . \quad (6')$$

The solutions of this equation depend upon the relative values of a, b, and k.

Case I.—If $b>\frac{a^2}{4}$, then the solution is

$$V_c=He^{-\frac{a}{2}t}\cos(pt-\delta)+\frac{R_2}{R_1}V_0\frac{1}{CL_2b}+\frac{1}{CL_2}\left[\frac{Le^{-kt}}{k^2-ak+b}+\frac{Me^{kt}}{k^2+ak+b}\right], \quad . \quad . \quad . \quad (7)$$

where H and δ are arbitrary constants. From the given initial conditions, namely,

$$V_c=V_G \quad \text{and} \quad i_2=0 \quad \text{at} \quad t=0,$$

these arbitrary constants can be determined, and i_2 may be deduced.

$$i_2=\frac{V_0-V_c}{R_1}-C\frac{dV_c}{dt}.$$

Hence the turning-point of i_2 may be found, being, indeed, given by the value of t for which

$$\frac{d^2V_c}{dt^2}+\frac{1}{R_1C}\frac{dV_c}{dt}=0.$$

This last equation is of the form

$$Ae^{-\frac{a}{2}t}\cos(pt-\phi)+Be^{-kt}+Ce^{kt}=0, \quad . \quad . \quad . \quad (8)$$

where A, B, C, ϕ, may all be functions of k. Theoretically it is possible to solve equations (8) and (5′) simultaneously

for k and T, so that the maximum value of i_2 may occur at the same instant as the minimum value of V_s, but in practice the process is very difficult, if not impossible. Assuming the calculation performed, there will be numerous real values of k corresponding to an equal number of values of T. It is in the choice of the particular value of k that the nature of the characteristic depends.

In Case I. let it be assumed that

$$0<T<\frac{2\pi}{p}.$$

Then i_2 will have only one turning-point in the range in which V_s has only one turning-point; and the corrresponding value of k will be the largest root of the equations (5′) and (8). Let this value be K. Then the first and most important condition that a characteristic of the type of fig. 5 may be obtained has been fulfilled. The second condition is that the characteristic should consist over part of its range of two almost superimposing curves of the general shape shown, approximately parabolic, elliptic or hyperbolic; it is impossible to say from a casual examination of the oscillograms. To show that this condition is satisfied by substituting for V_s the expression $Le^{-kt}+Me^{kt}$, it is unfortunately necessary to argue from the particular to the general; the examination of the general case is beyond the mathematical ingenuity of the writer. However, although laborious in the extreme, the numerical calculation of i_2 and V_s for particular cases leaves little doubt as to the essential validity of the theory.

Example.—In one of the writer's circuits the resistances, inductance, and capacity were as follows:—

$R_1=200$ ohms, $R_2=2$ ohms, $L_2=30$ microhenries, $C=2$ microfarads, $V_0=250$ volts. The oscillogram for this circuit gave $V_G=110$ volts, $V_A=20$ volts, so that

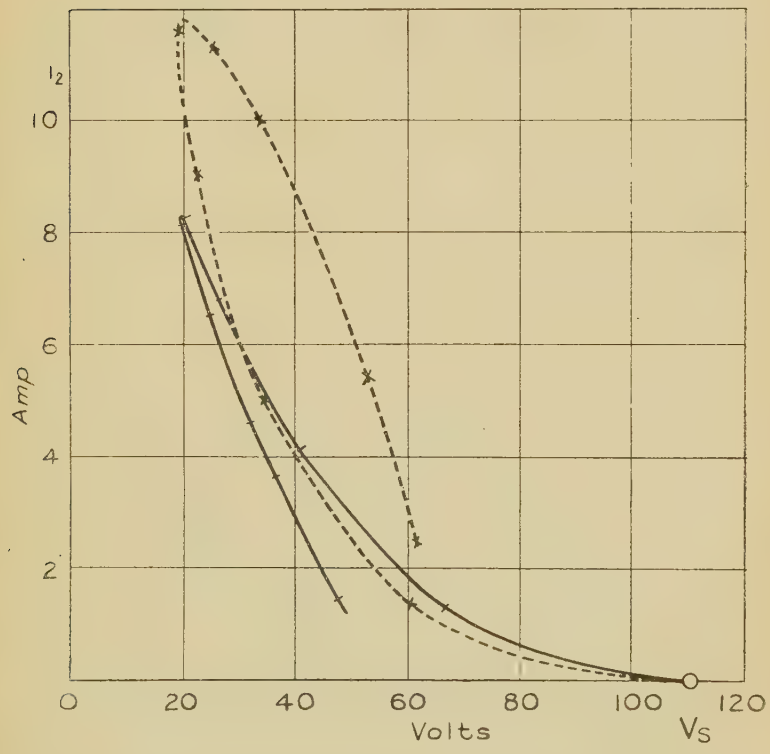
$$V_s=109{\cdot}09e^{-kt}+0{\cdot}91e^{kt}.$$

By the method of trial and error a value of k was chosen, $k=1{\cdot}3p$ where p is the pulsatance of the circuit. The complete differential equation for the circuit then became

$$\frac{d^2V_c}{dt^2}+6{\cdot}825\,.\,10^4\frac{dV_c}{dt}+1{\cdot}66\,.\,10^{10}V_c=4{\cdot}11\,.\,10^{10}$$
$$+1{\cdot}644\,.\,10^{10}[109{\cdot}1e^{-9{\cdot}147}\,.\,10^4t+0{\cdot}91e^{9{\cdot}147}\,.\,10^4t].$$

This equation was solved, and i_2 was calculated. Then from the numerical values of V_s and i_2 the graph shown by the full line in fig. 7 was drawn. A comparison of this curve with the characteristic of fig. 6 (Pl. XVI.) shows that they are very similar. The oscillogram was magnified and compared directly with a graph drawn to scale.

Fig. 7.



It was found that the rising part of the theoretical curve was indistinguishable from the experimental characteristic.

However, as is evident, the falling part of the characteristic is *not* in good agreement with the facts. This is due to the rather crude method of choosing k. The dotted curve in fig. 7 was obtained by means of a similar calculation, using a higher value for k. It is clear that a value of k could be chosen between the two which would

give the experimental curve. This value of k would lie rather nearer to $1{\cdot}3p$ than to the other value.

Hence it may be concluded that it is possible, in certain cases at least, to represent with some degree of accuracy the potential across an oscillating arc by an expression of the type shown in equation (5).

Fig. 7 (Pl. XVI.) shows what *may be* a limiting case of this régime. The potential across the arc in this case falls practically to zero before beginning to rise again. This means that the ratio L/M is exceedingly large, and therefore, if the minimum value of V_s occurs in the neighbourhood of $t=\frac{\pi}{2p}$, k must be very large also. An examination of equation (7) indicates that when k is very large the aperiodic exponential part of V_c' (the particular integral minus a constant) becomes small, and therefore i_2 takes the simple form

$$i_2=\mathrm{R}+\mathrm{S}e^{-\frac{a}{2}t}\sin(pt-\phi)$$

where ϕ is small.

Calculation for a particular case again shows that this equation for i_2 gives the required characteristic the asymmetry of the traces in fig. 7 (Pl. XVI.) being clearly due to the asymmetry in the charging of the condenser. Of course, as in the case of the simple condenser "spark," the potential V_s has to change sign when the arc goes out, a high-frequency oscillation bringing about the conditions necessary for the starting of the glow with a negative potential across the gap. Incidentally, it is important to observe that where the gap rectifies the current (as in the case of carbon-copper electrodes) the mechanism is similar, but the high-frequency oscillation (more or less modified by thermionic conduction across the gap and the charging current to the condenser) makes at least three-quarters of a complete cycle before the discharge recommences.

Case II.—Let it again be assumed that $b>\frac{a^2}{4}$ in equation (6′), but in this case when equations (5′) and (8) are solved for k, let a value of the latter be chosen for which $\frac{2n\pi}{p}<\mathrm{T}$ and n is large—of the order 10 perhaps.

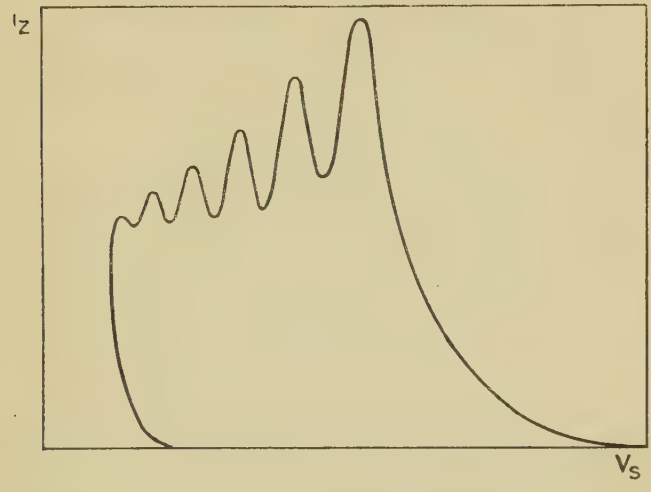
This is equivalent to choosing a small value of k, or, putting the matter in another way, it is equivalent to assuming a high frequency in the oscillating circuit. Then the current i_2 has several maxima and minima before V_s reaches its minimum, and the i_2, V_s characteristic takes an entirely different form.

As before,

$$i_2 = c_1 + c_2 e^{-\frac{a}{2}t} \sin(pt - \phi) + c_3 e^{-kt} + c_4 e \ ,$$

but in this case, since k is small compared with p, the potential across the discharge falls much more slowly,

Fig. 8.



and a simple calculation shows that c_2 is small compared with either or both of the coefficients c_3 and c_4. Without entering into the laborious calculation it may simply be stated that an i_2, V_s characteristic of the form shown in fig. 8 is obtained, with, of course, minor variations depending upon the circuital constants. This type of characteristic has been observed by the writer when the inductance in the discharge circuit is small, the charging current being large, and it is under these conditions that the theory given might be expected to apply. For, if the inductance L_2 is small, and the resistance R_2 is also small, there can be little difference between the potential across the arc and across the condenser. In

other words, energy is more easily supplied by the condenser to the discharge. This decreases the value of k and so changes the régime to give a characteristic of the form shown. The writer has also observed this form of characteristic using an induction coil as the source of potential with a very small capacity in parallel with the gap. It is, in fact, the characteristic of the "pulsating arc" familiar to all users of induction coils.

It is now clear that there is an alternative explanation of fig. 7 (Pl. XVI.). It is not at all certain from an examination of the photograph that the current does not oscillate several times in the high-current part of each trace. Hence it is just possible that the correct interpretation of this figure is along the lines of the discussion of Case II. Again making the ratio L/M large, it is still possible to keep k small by increasing T, and this method of calculation ensures that V_s will not decrease or increase very rapidly in the neighbourhood of the turning-point. Hence, while V_s remains practically constant, i_2 may vary considerably. This sequence of events fits in very naturally with the appearance of fig. 7 (Pl. XVI.), but without further experiment it is impossible to say which is the correct explanation. In any event the point is of minor importance. Both régimes are possible in practice; the value of the inductance L_2 determines which of them will occur.

Case III.—The question must now be asked—if $a^2/4$ is greater than b, so that the condenser inductance circuit will not oscillate with the gap closed, is it possible that the inclusion of an air-gap in which a discharge is flowing can produce oscillation? At first sight it would appear that the circuit of fig. 6 might still have a period of oscillation, that period being determined by the time taken to charge the condenser through the resistance R_1; but, as will be proved immediately by the *reductio ad absurdum* method, this is impossible in practice.

If $a^2/4>b$,

$$V_c=He^{-\frac{a}{2}t}\cosh(pt-\delta)+\frac{R_2}{R_1}V_0\frac{1}{CL_2b}+\frac{1}{CL_2}\left[\frac{Le^{-kt}}{k^2-ak+b}+\frac{Me^{kt}}{k^2+ak+b}\right].$$

(1) If k is small compared with a and p, then after a negligibly small time

$$i_2 \doteqdot A - B \sinh (kt)$$

$$V_s \doteqdot C + D \cosh (kt)$$

and, if these approximations are to hold, i_2 must be negative or the variable part must be very small. If i_2 is negative, the theory breaks down, since the current cannot pass from a positive to a negative value without the cessation of the discharge. If, on the other hand, the variable part of i_2 is small, the discharge characteristic will not be of the required form. In fact i_2 will be approximately constant. This amounts to a stabilization of the discharge. The arc ceases to hiss or sing, and the action of the condenser may be neglected.

(2) If k is large compared with p,

$$i_2 = A' - B' e^{-\frac{at}{2}} \sinh (pt - \phi),$$

and this expression is open to exactly the same objection as the last. It is peculiar that this case does not give an oscillating arc, if only because the engineer's formula for the arc characteristic (equation (4)) may be obtained by combining functions of the form

$$i_2 = A + B \sinh (st),$$
$$V_s = C + D \cosh (st).$$

Yet, so far as the writer can see, it is not possible to derive these expressions in the proper form from the circuital equations.

Conclusions.—It has been shown that by assuming $V_s = Le^{-kt} + Me^{kt}$ as the time-equation for the potential across a discharge, the typical oscillating discharge characteristics may be obtained.

(*a*) In particular the normal characteristic of a hissing or singing arc has been obtained, and it has been suggested that the arc characteristic of a capacity discharge is merely a limiting case of this phenomenon, where the maximum current across the gap may be maintained at a very low potential.

(*b*) In another case it has been shown that the well-known "pulsating-arc" characteristic can be obtained using the same hypothesis, by presupposing a value of k consistent with the experimental conditions.

General Remarks.

If the theories given above are correct, a very important problem emerges from the discussion. Why does the potential across the gap vary exponentially with the time ? Already the writer has several views upon the subject, but since they lack both precision and verification they will be withheld from present publication. There can be no doubt that a close connexion exists between the value of k and the circuital constants : also, if the gas pressure or gap length is varied, a variation in k will undoubtedly result. It is suggested that there is here a fruitful field for investigation which will undoubtedly lead to a better understanding of discharge phenomena.

Acknowledgments.

I again wish to thank Professor Taylor Jones for the great interest he has taken in the work, and for his valuable encouragement. Professor Milner called my attention to the problem, and I am very grateful to him for his kindness. I am also much indebted to my wife, without whose valuable assistance the rather laborious calculations would have taken double the time to perform. The experimental work was performed in the Research Laboratories of the Natural Philosophy Department of the University of Glasgow.

References.

(1) Phil. Mag. xiii. p. 824 (1932).
(2) Phil. Mag. November 1912.
(3) *Ann. d. Phys.* xxxvi. p. 281 (1911).

December 1932.

LVI. *Some Elementary Considerations connected with Modern Physics.*—I. *The Larmor-Lorentz Transformation. By* Sir Oliver J. Lodge, *F.R.S., and* Professor Alfred Lodge, *M.A.**

LITERARY people, besides pursuing their discoveries and producing their masterpieces, sometimes play about with their subject as if they enjoyed it; and thereby they consolidate and make palatable and intelligible what

* Communicated by Sir Oliver Lodge.

might otherwise be rather arid and esoteric. Literary and artistic criticism have become an occupation in themselves, and commentaries serve a useful purpose. There is very little play in modern physics. The subject is difficult, the ground is new, and workers are keen, making up for their fewness by concentration and industry. Repetition and gleaning seem comparatively frivolous while the main crop is waiting to be garnered in, and new adventures are calling in every direction. Yet a time for consolidation and digestion must come, and novelties require to be assimilated as well as discovered.

With this preamble we will deal with a few elementary matters.

Everyone must have noticed that the composition of relative velocities has the same form as the composition of a trigonometrical function of two angles, especially of imaginary angles; and sometimes trigonometrical expressions are neater than algebraic ones. Now Relativity insists first of all that there is in nature a constant absolute velocity c, a characteristic of space or of the medium which fills space, which turns out to be of the same order, and perhaps of the same value, as the undoubted velocity with which space transmits periodic transverse disturbances. Furthermore, whatever may happen to waves under exceptional circumstances such as those in electrified space, the velocity of a particle or of a quantum of energy can never exceed c, and indeed can only approach it asymptotically; so that a particle's velocity bears to c a ratio which may approach, but cannot exceed, unity. (While at the same time that factor of the particle's motion which confers on it momentum tends to increase without limit.)

A trigonometrical function with this property is well known, namely, the hyperbolic tangent of an angle; or, what is the same thing, the imaginary circular tangent of an imaginary angle. So let us represent $\frac{u}{c}$ as $\tanh\theta$, and see what happens to the momentum factor m_0, and to the familiar Larmor-Lorentz transformation.

The FitzGerald contraction $\frac{1}{\beta}$ or $\sqrt{1-\frac{u^2}{c^2}}$ becomes $\sqrt{1-\tanh^2\theta}$, that is sech θ. Hence β is simply cosh θ, a function which is readily pictured as the curve of a

hanging chain. So that β rises from unity when $\frac{u}{c}$ is small, as it usually is, and tends towards infinity as $\frac{u}{c}$ approaches unity.

Accordingly, taking an auxiliary angle for subsequent interpretation, and putting $\beta=\cosh\theta$, and $u=c\tanh\theta$, the conditions are satisfied; and, moreover, the transformation of the coordinates of a particle moving with velocity u, or of a particle referred to moving axes, is :—

$$\left.\begin{aligned} x'=\beta(x-ut)&=x\cosh\theta-ct\sinh\theta,\\ \text{and}\qquad ct'=\beta(ct-ux/c)&=ct\cosh\theta-x\sinh\theta,\end{aligned}\right\} \quad . \quad (1)$$

which equations are identical with the old established form for transition from one set of axes to another set rotated through the angle $i\theta$, namely :—

$$\left.\begin{aligned} x'&=x\cos i\theta+y\sin i\theta,\\ y'&=y\cos i\theta-x\sin i\theta,\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (2)$$

where

$$y=ict, \quad \text{and} \quad y'=ict'.$$

This identity at once suggests that the Larmor-Lorentz transformation to moving axes is equivalent to a rotation of the axes about an imaginary angle $i\theta$. We do not properly know what relative locomotion such as u really is; absolute motion through the ether may very well involve something equivalent to rotation.

The problem is whether any physical significance may be attached to this rotation. It is known that time can be converted into space by multiplying it by the factor ic. And if there is a constitutional rotation in the ether it may very well be represented by an imaginary angle $i\theta$, since it is clearly not an ordinary revolution in space, and its real physical interpretation is not expressible in space relations without the converting factor i. Time cannot be real space: ict is only space of a fourth dimension, the factor i being appropriate for expressing a dimension at right angles to the three simple space dimensions. It represents rotation through a right-angle, and thus enables us to pass from three to four dimensions and to include time.

But ordinary rotation is an affection of matter. What rotation can mean in ether we do not know; it presumably

has an analogy with material rotation, and might reasonably be prefixed by a $\sqrt{-1}$. This mode of separating ether terms from matter terms may be serviceable, for it will keep them from being mixed up. A decay of matter e^{-nt} may impart to the ether an oscillation e^{-int}. The rate of decay would thus determine the frequency of the consequent radiation. That is to say, a disappearance of matter at the rate h might give rise to radiation $h\nu$.

To get the new or transformed velocity appropriate to a constant u when the old velocity $\frac{dx}{dt}$ is v, that is to compound v and u, we must reckon the value of $\frac{dx'}{dt'}$; and that comes out thus :—We can take v as $c \tanh \phi$, another auxiliary angle, and since by (1)

$$dx' = \cosh\theta\, dx - c \sinh\theta\, dt$$

and

$$c\,dt' = c\cosh\theta\, dt - \sinh\theta\, dx,$$

it follows that

$$\frac{1}{c}\cdot\frac{dx'}{dt'} = \frac{v\cosh\theta - c\sinh\theta}{c\cosh\theta - v\sinh\theta}$$

$$= \frac{\tanh\phi\,.\cosh\theta - \sinh\theta}{\cosh\theta - \tanh\phi\sinh\theta}$$

$$= \frac{\sinh\phi\cosh\theta - \cosh\phi\sinh\theta}{\cosh\phi\cosh\theta - \sinh\phi\sinh\theta}$$

$$= \frac{\sinh(\phi-\theta)}{\cosh(\phi-\theta)} = \tanh(\phi-\theta).$$

So to summarize:

$$\frac{dx}{dt} = c\tanh\phi$$

and

$$\frac{dx'}{dt'} = c\tanh(\phi-\theta),$$

where

$$u = c\tanh\theta.$$

The Larmor-Lorentz Transformation—Graphically Treated.

Let us begin again.

Imagine two observers, one of whom moves relatively to the other with a positive velocity u along a direction

which we shall choose as the axis of x, and let them observe an interval between two events separated both in space and time ; and let x, t be the estimates of distance (along the x-axis) and of time, made by one observer, while x', t' are the corresponding estimates made by the observer who is moving with the above velocity u. Then the equations connecting these two pairs of quantities are known to be

$$x'=\beta(x-ut), \quad t'=\beta\left(t-\frac{ux}{c^2}\right),$$

$$x=\beta(x'+ut'), \quad t=\beta\left(t'+\frac{ux'}{c^2}\right),$$

where c denotes an ultimate velocity unsurpassed by any material thing, and β is a coefficient required to make the second pair of equations consistent with the first pair, and therefore with the principle of relativity.

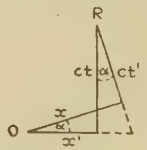
To find β, substitute in the second pair the values of x' and t' from the first pair. Then

$$x=\beta^2\left(x-ut+ut-\frac{u^2x}{c^2}\right)=\beta^2x\left(1-\frac{u^2}{c^2}\right)$$

and

$$t=\beta^2\left(t-\frac{ux}{c^2}+\frac{ux}{c^2}-\frac{u^2t}{c^2}\right)=\beta^2t\left(1-\frac{u^2}{c^2}\right).$$

Each of these identities requires β to be equal to $\left(1-\frac{u^2}{c^2}\right)^{-\frac{1}{2}}$. Therefore $\frac{1}{\beta}$ is the FitzGerald contraction.



To represent these graphically, let $u=c\sin\alpha$, whence $\beta=\sec\alpha$.

$$\therefore \quad \left.\begin{aligned} x'&=x\sec\alpha-ct\tan\alpha,\\ ct'&=ct\sec\alpha-x\tan\alpha,\end{aligned}\right\}$$

and

$$\left.\begin{aligned} x&=x'\sec\alpha+ct'\tan\alpha,\\ ct&=ct'\sec\alpha+x'\tan\alpha,\end{aligned}\right\}$$

equations which are represented by the accompanying figure.

These equations satisfy the equation $(ct)^2-x^2=(ct')^2-x'^2$, which is independent of α, and is therefore the same for different values of the relative speed u of the two observers, and therefore invariant for all observers. If, therefore, x and ct are represented in a diagram as the rectangular coordinates of a point P, and x' and ct' are the coordinates of a point P', this invariance indicates that P and P' lie on the same hyperbola for all values of α.

To find x', ct' when x, ct, and also the angle α, are given.

Take one event at the origin ,where we may suppose the observers to be located, one at rest, and the other having relative velocity u along the axis of x ; and let the second event be represented (to the first observer) by a point P (fig. 1) whose coordinates are (x, ct). The problem is to find x' and ct' for the given value of $\alpha=\sin^{-1}\frac{u}{c}$. With centre O, radius ct, describe a circle, and draw the radius OT at the angle α with the y axis. Draw the tangent TQ cutting PN the ordinate of P in Q. Then

$$QN=ct', \quad \text{and} \quad TQ=x'.$$

For, if QT cuts the y-axis at R, and QM is parallel to ON, it is clear that

$$\begin{aligned} TQ&=RQ-RT \\ &=x\sec\alpha-ct\tan\alpha \end{aligned}$$

and

$$\begin{aligned} NQ&=OR-RM \\ &=ct\sec\alpha-x\tan\alpha. \end{aligned}$$

Also it is clear that

$$\begin{aligned} x(=MQ)&=TQ\sec\alpha+OM\tan\alpha \\ &=x'\sec\alpha+ct'\tan\alpha, \end{aligned}$$

and

$$\begin{aligned} ct(=OT)&=OM\sec\alpha+TQ\tan\alpha \\ &=ct'\sec\alpha+x'\tan\alpha. \end{aligned}$$

And $$OQ^2=(ct)^2+x'^2=(ct')^2+x^2,$$

so that $$(ct)^2-x^2=(ct')^2-x'^2.$$

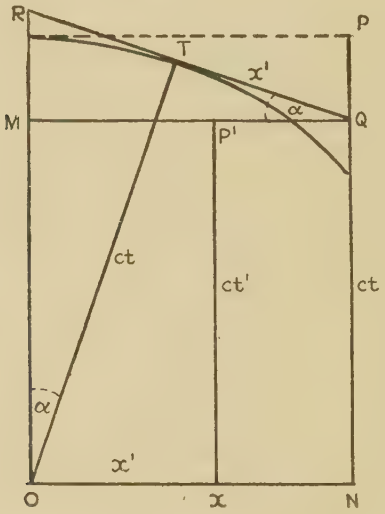
If now MP' is taken equal to TQ, on the line MQ, the coordinates of P' are (x', ct').

So P' represents the occurrence of the event P as it

appears to an observer at O who is moving along the x axis with a speed $u=c \cdot \sin \alpha$.

The value of the invariant quantity $(ct)^2-x^2$ is positive, zero, or negative according as $ct>$, $=$, or $<x$. Its geometrical significance in each case can be illustrated by taking three points A, B, C having the same value of ct, but with values of x respectively less than, equal to, and greater than ct. We will find the corresponding values A′, B′, C′ for a given value of α, and then consider their loci as α changes. The diagram (fig. 2) also shows

Fig. 1.



the parallelogram into which the rectangle OYCc is distorted.

The sides of this parallelogram are inclined to the corresponding sides of the rectangle at an angle γ such that $\tan \gamma = \sin \alpha$. For consider the lines AC, A′C′, and let the coordinates of A and C be (x_1, ct) and (x_2, ct), the value of t being the same for both A and C, but the values of x being different ; then, since

$$x' = x \sec \alpha - ct \tan \alpha ;$$
$$ct' = ct \sec \alpha - x \tan \alpha ;$$

$$\left.\begin{aligned} \therefore \quad x_1' - x_2' &= (x_1 - x_2) \sec \alpha \\ \therefore \quad ct_1' - ct_2' &= (x_2 - x_1) \tan \alpha \end{aligned}\right\} \text{as } ct \text{ is constant};$$

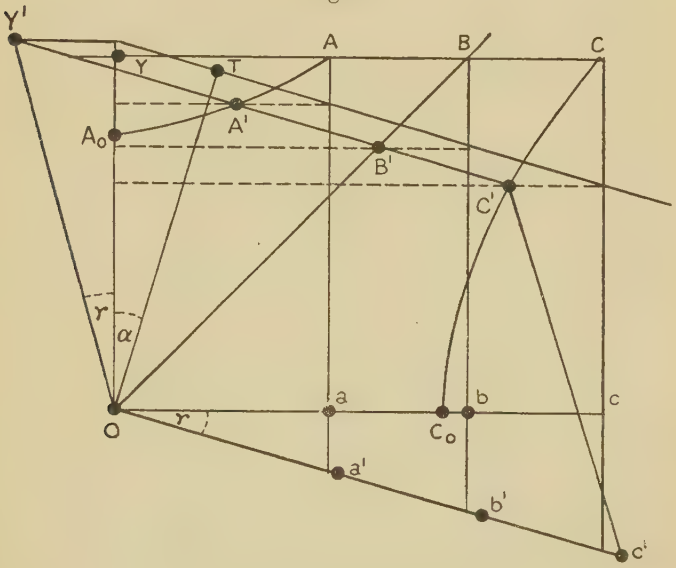
hence, if γ is the angle between AC and A′C′, which is the same as the angle between Oa and Oa',

$$\tan\gamma = \frac{\tan\alpha}{\sec\alpha} = \sin\alpha = u/c.$$

Similarly, the angle YOY′ is also equal to γ.

So, in general, the derived or projected figure is distorted angularly as well as in length, but the ratio of A′B′ : B′C′ is equal to the ratio of AB : BC, and parallel lines remain parallel.

Fig. 2.



The points AA′ lie on a hyperbola whose semi-transverse axis is OA_0, A_0 being the point where the circle, centre a, radius aA, cuts the y axis. This is the locus of A′ as α varies. Similarly, the locus of C′ is the hyperbola having OC_0 as its semi-transverse axis, C_0 being the point where the circle, centre Y, radius YC, cuts the x axis.

The 45° line, containing B and B′, is an asymptote of the system of hyperbolæ, and the other asymptote is perpendicular to it.

There are two directions in which there is no angular distortion, viz., if lines are drawn parallel to the asymptotes their representatives in the system, or, as we may

call them, their projections, are also parallel to the asymptotes, but they will be altered in length.

For

$$x' = x \sec \alpha - ct \tan \alpha,$$

$$ct' = ct \sec \alpha - x \tan \alpha$$

$$\therefore \quad x' + ct' = (x + ct)(\sec \alpha - \tan \alpha)$$

and $$x' - ct' = (x - ct)(\sec \alpha + \tan \alpha).$$

Therefore, if $x+ct$ is constant, $x'+ct'$ will also be constant, but a smaller constant ; and if $x-ct$ is constant, $x'-ct'$ will be a bigger constant.

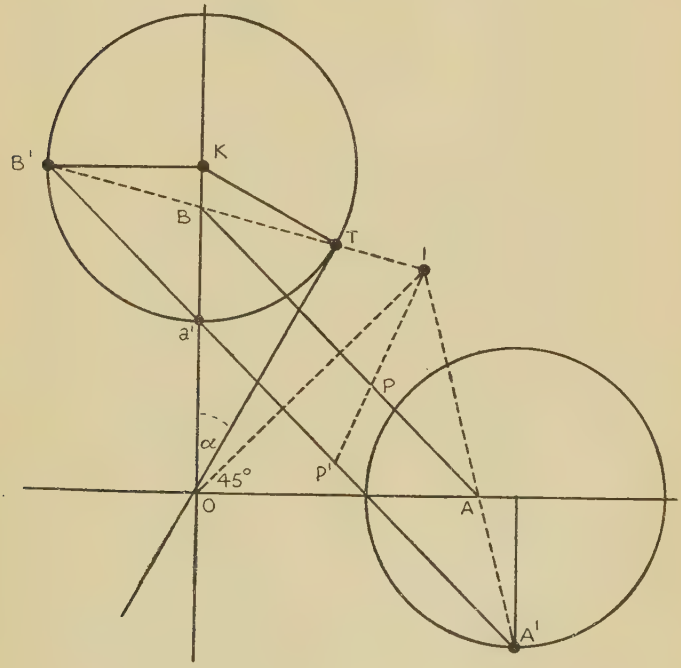
The importance of these lines lies in the fact that they are the time-distance tracks of light travelling from an event towards the observers, assuming that the light is travelling with the velocity c. The lines $x+ct=$constant emanate from points on the right of the diagram, when x is positive, and the lines $x-ct=$constant come from points on the left. We shall consider only points on the right, utilizing the lines $x+ct=$constant. It is important to see that the light-tracks from (x', ct') are parallel to those from (x, ct) in accordance with the condition that light travels with the same velocity c whatever may be the relative speed u.

The accompanying diagram (fig. 3) shows two corresponding tracks, AB and A′B′, the one starting from A or from any event P on the line AB, and the other starting from A′ or from a point P′ corresponding to P. The dotted lines in the figure show one method of finding P′ when P is given. This construction is based on the fact that P′ must divide A′B′ in the same ratio as P divides AB. The distance OB shows the time at which the observer will see the event at A if $u=0$, and Oa' shows the time at which he will see the same event if the relative velocity is $u=c \sin \alpha$. In each case OB is the observer's time track : the diagram is so adjusted that the velocity u is given to the object on which the event occurs. This will be alluded to again and exemplified in the sequel.

Although an attempt is made in these diagrams to depict time as if it were space, remember that a time-diagram cannot be static. It inexorably progresses, turning the future into the past. Consequently the line through the observer has to move upwards : or else the whole diagram has to move downwards past O. Events

are not seen when they actually occur; an observer has to wait till the light from them reaches him. The velocity of light in these diagrams is represented by the slope of 45°, since the distance travelled by the light in time t is equal to ct, so that $\frac{dx}{c\,dt} = \pm 1$ according to its direction of travel. The light-line, therefore, from an event P is a line of 45° towards the y axis, *i. e.*, towards

Fig. 3.



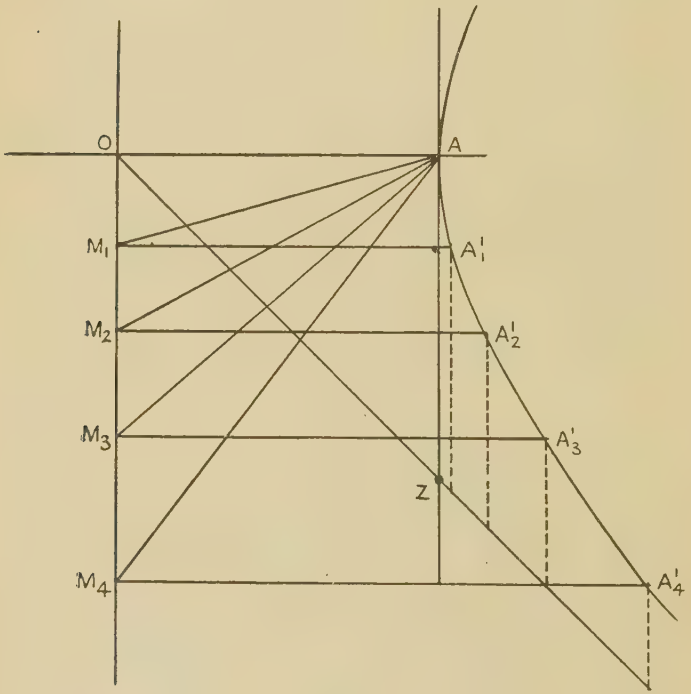
the line $x=0$ along which the observer is travelling in time.

Consequently an event occurring at P (x, ct) is seen at time T given by $cT=x+ct$ by the stationary observer, and is seen by the other observer at time T′ given by the equation $cT'=x'+ct'$. Note that the moving observer, considering himself at rest, has placed P′ in his diagram so that the axes of x' and t' coincide with those of x and t.

Scale of Diagrams.—We have assumed the scales of x and ct to be equal, so that when $x=ct$ the line is at 45°.

As these diagrams are most useful when applied on a cosmic scale, it will be better really to measure x and ct both on the light-time scale, *i. e.*, really to measure x by the time light would take to travel the distance, so that we really depict $(x/c, t)$ rather than (x, ct). It is merely, then, a matter of scale : in either case the light-tracks are at 45° in the map.

Fig. 4.

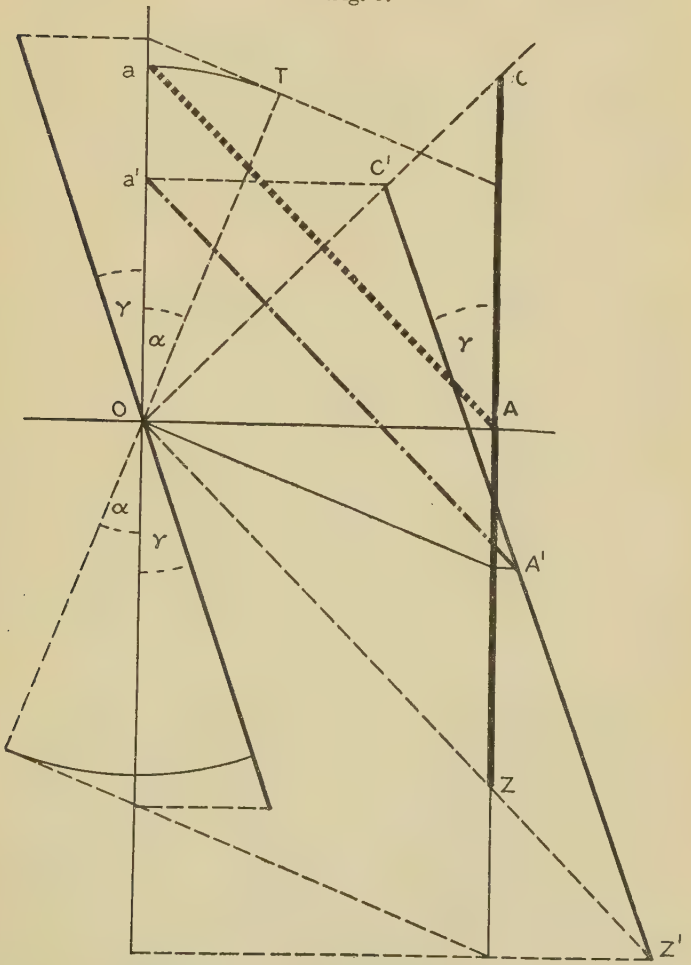


Application to Astronomy.

The geometry indicated in the preceding pages can be utilized for the purpose of what may be called a celestial time-distance map of, for instance, the planets as seen from the earth. Take, for instance, the planet Jupiter and an event on it such as an eclipse of one of its satellites. Instead of considering two observers, one at a fixed distance from Jupiter and the other moving towards it with velocity u, take one observer, and

consider the time-distance track of Jupiter according as it is at a stationary distance or is approaching the observer with velocity u, and on this track show where and when the eclipse would appear to take place in the

Fig. 5.



two cases. There are two distinct times in each case: (1) when the eclipse *occurs*, (2) when it is *seen*. The eclipse occurs at time t' or t according as there is or is not relative motion; the eclipse is seen when $T' \equiv t' + \frac{x'}{c}$ or $T \equiv t + \frac{x}{c}$ under the same variety of conditions.

In fig. 5 the time-tracks of Jupiter are ZAC for the case when its distance from the observer is constant, *i. e.*, when $u=0$, and Z′A′C′ for the case when it is moving towards the observer with velocity u. The angle that this latter track makes with AC is γ, where $\tan\gamma=\frac{u}{c}$. This angle is the same as would be given for such a track in any Euclidean distance-time chart, its gradient being proportional to its velocity : the only peculiarity is its displacement in space, as it is not drawn from A as in an ordinary track such as in railway time-tables, but from A′, a point earlier in time than A, and further away from the observer, but so placed that the track A′C′ crosses the line OA in front of A.

If the eclipse occurs at A when the distance x remains constant, the diagram puts it at A′ if Jupiter is approaching O with velocity u. This enables the observer in this latter case to see the eclipse earlier, viz., after an interval Oa' instead of Oa, although the light travels at the same rate in either case : the light-tracks Aa, A′a' are parallel, but not coincident.

Note.—It is not necessary actually to draw these light-tracks to find the time-intervals : they can equally be reckoned from the distances of A, A′ from the 45° line OZZ′, measured in the time direction, since these are parallel and equal to Oa, Oa' respectively. Fig. 4 shows this for different observers, or for one observer with different speeds.

We have shown that the estimated time-intervals between the *occurrence* of an event and its being *seen* are indicated by its time-distance from the asymptote through the observer ; and that these time-distances, shown in the diagram by AZ and the parallel dotted lines from the position of A′ (a position which is always on the hyperbola whose vertex is at A), are less and less according as the relative velocity u becomes greater and greater.

So, if there is a second event occuring later (not shown) its estimated time-interval will also be proportional to these distances. It does not matter whether the distance of Jupiter remains the same : it will not if it is really approaching : the figure will be *similar* in all respects to the one drawn.

So, the time-interval between two events as estimated by a moving observer will depend on the value of u, and

will be proportional to the distances of the various A′ points from the asymptote. As u increases, these intervals become less and less, as is rather distinctly shown by this hyperbola diagram (fig. 4).

[The mode of finding these points is shown by the construction already stated, viz., making MA′=MA= OA sec α, where α is OAM.]

We are taking the eclipse as occurring at time $t=0$, *i. e.*, "now" to one observer; it is the most interesting case, but in all cases the intervals between occurring and being seen obey similar laws to those indicated above, for all events are "now" at some time or another!

Recapitulation, with special reference to Fig. 5.

Let (x, ct) denote the distance away, and the time of occurrence, of an event, when the distance of the observer from the place where the event occurs is constant. It does not follow that the observer sees it then, since light takes time to travel the distance. In the diagram he has to wait till the light-line arrives at $x=0$. So all events are seen in the future if they lie above the $-45°$ line OZ. Taking the line OX as the "now" line, the line of zero time, events below this line have already occurred, but are *seen* in the future. So OZ is the line of simultaneous visibility, though OX is the line of simultaneous occurrence.

Now if the event is occurring at a place which is approaching the observer with velocity u, its occurrence appears to be earlier and at a slightly greater distance, but any event which would have been seen in the future when $u=0$ will still be seen in the future, though at an earlier moment; since any event A, which lies above OZ when $u=0$, will still lie above it whatever u may be, as long as u is less than c.

So far we have considered that, when Jupiter approaches the observer, an event appears to occur earlier than if it were at relative rest. In the diagram (fig. 5) the time axis is fixed, and the path of Jupiter is tilted instead. This tilt is $\gamma=\tan^{-1} u/c$. The apparently earlier occurrence is certainly indicated by the diagram and seems to be justified by the equations. But in what follows we endeavour to show that this curious time distortion, so repugnant to common sense, need not necessarily be considered to be a physical fact, though it still remains a

useful mathematical fiction. If we can permit ourselves to imagine that light may reach the eye of an approaching spectator with a speed $c+u$, we shall see that it is possible to consider the time of occurrence as perfectly definite whatever the relative motion may be.

Seeing an Event occurring now.

To reckon the times, T and T′, of seeing an event after its occurrence, i. e., *after the time when* $t=0$.

If the object on which the event occurs is at a constant distance x from the observer, then, when $t=0$, the time required before the event is seen is

$$\mathrm{T}=\frac{x}{c}.$$

But if the observer is approaching the object with velocity u, i. e., if

$$\frac{dx'}{dt'}=-u,$$

the time at which the event will be seen will be T′, where

$$\mathrm{T}'=t'+\frac{x'}{c}.$$

When $t=0$, the equations

$$x'=x\sec\alpha-ct\tan\alpha,$$
$$ct'=ct\sec\alpha-x\tan\alpha,$$

become

$$x'=x\sec\alpha,\quad ct'=-x\tan\alpha;$$

$$\therefore\quad \mathrm{T}'=\frac{x}{c}(\sec\alpha-\tan\alpha).$$
$$=\frac{x}{c}\cdot\frac{1}{\sec\alpha+\tan\alpha}$$
$$=\frac{x}{c}\cdot\frac{\cos\alpha}{1+\sin\alpha}$$
$$=\frac{x\cos\alpha}{c+u}$$

This expression is actually less than $\frac{x}{c+u}$, which would have been its Euclidean value under a joint speed of

c and u. It *is* this fraction, with $x\cos\alpha$ instead of x, i. e., with the FitzGerald contraction applied to x. (To a first approximation the two values are equal, since $1-\cos\alpha$ is of the second order of small quantities compared with $\sin\alpha$.)

Referring to fig. 6, if the event A occurs on an object (*e. g.*, Jupiter) which is at a fixed distance from the observer at O, he will see the event at S, *i. e.*, when the light reaches S. If, however, the object is approaching the observer with velocity u, he will see the event at S′, *i. e.*, when the light reaches S′.

A′ is a point on the rectangular hyperbola of which A is the vertex, and the space-time track of the moving object, viz., A′BC′ is the tangent at A′ to the hyperbola, and slopes at an angle γ with the time axis OS, where $u=c\tan\gamma=c\sin\alpha$.

If this tangent cuts OS at T, the line TA makes an angle α with OS, and is perpendicular to MA, where

$$\mathrm{MA}=\mathrm{MA}'=x'=x\sec\alpha,$$

and also

$$\mathrm{OA}=x,\quad t=0,$$

$$\mathrm{OB}=x\cos\alpha$$

$$\mathrm{OS}'=x\sec\alpha-x\tan\alpha$$

$$=c\cdot\frac{x\cos\alpha}{c+u}=c\mathrm{T}'.$$

S″ is the seeing point if the object is moving away with the same velocity, the positive time direction being downwards in this case, using the same diagram.

$$\mathrm{OS}''=x\sec\alpha+x\tan\alpha$$

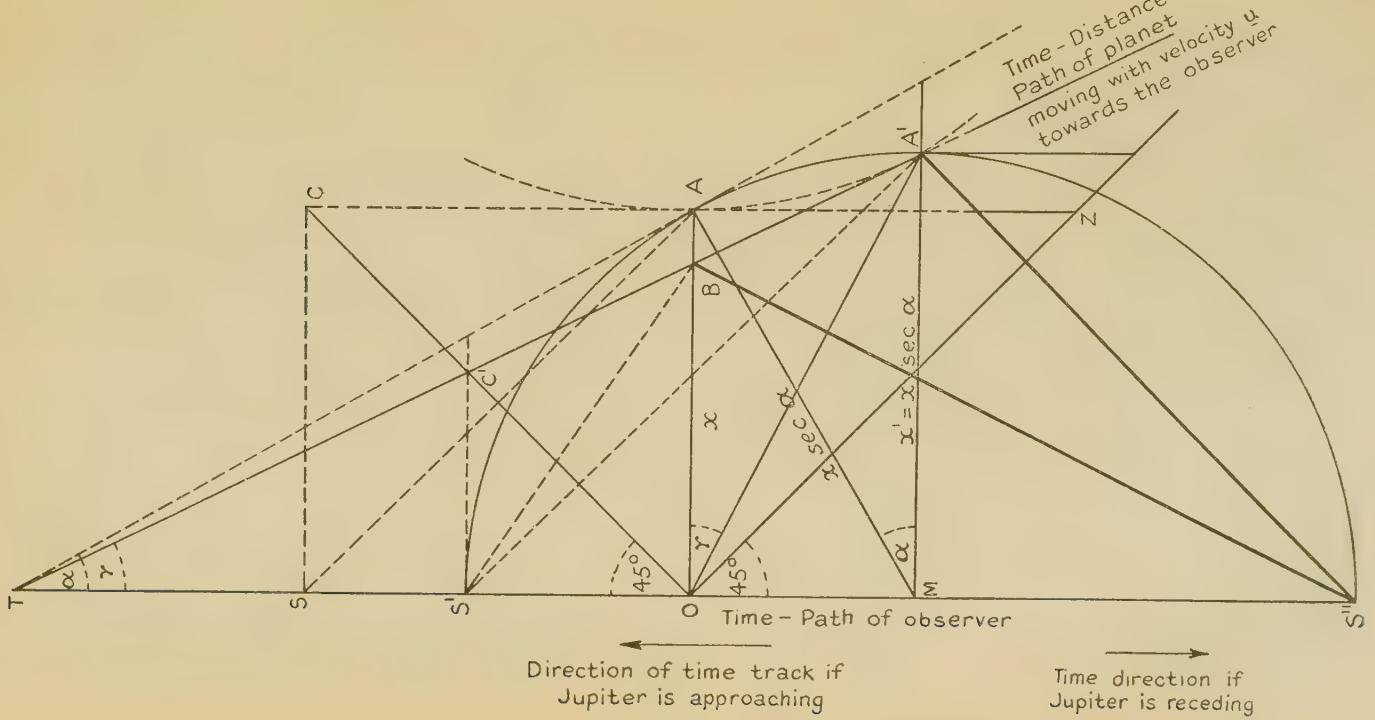
and

$$\mathrm{T}''=\frac{x\cos\alpha}{c-u}.$$

It is possible for all observers to agree when an observed event really occurred.

Now let us see what difference the relative velocity of the object and observer makes. An event which really occurs at A seems to occur either at A′ or at B; that is, it seems to occur earlier and a little further away than it actually did, or else time remains unaltered, but distance is foreshortened; that is, we interfere either

Fig. 6.


Time - Distance
Path of planet
moving with velocity u
towards the observer
A'
A
C
C'
B
Z
x
x sec α
x' = x sec α
α
γ
45°
45°
T
S
S'
O
M
S''
Time - Path of observer
Direction of time track if
Jupiter is approaching
Time direction if
Jupiter is receding

with time and distance or with distance only. But the interference with time introduces the artificiality of making the event occur before it does; while the disturbance of distance is only extending the FitzGerald contraction till it is of cosmic importance. A′ is best in a diagram, but philosophically B is better, since time is not interfered with, and distance is only altered in a way to which we are accustomed. No matter whether the event seems to occur at A′ or at B, it is seen at the same time. A′ is an inference, just as B is. But B has the remarkable advantage of being in accord with common sense. For the time taken before the event is seen by the moving observer is $\frac{x \cos \alpha}{c+u}$, that is, the velocity of the observer is added to the velocity of light, and as the two are approaching each other, this is what one would expect by common sense. The velocity of light is, of course, constant, but measured by a moving observer comes out $c \pm u$ according as the observer is approaching towards or receding from the source of light.

This is in agreement with common sense, but is not in accord with orthodox relativistic teaching. By it a *Doppler effect* is given to the moving observer, but the *velocity* as measured is supposed to be the same as if he were stationary, that is to say, orthodox teaching supports the point A′ instead of the point B, and makes the light take the 45° track from A′, in spite of the relative motion; whereas the point B has the advantage of lying on the same occurrence line as A, while the light-track is sloped at an angle corresponding to the combined velocities c and u.

We being the moving observers observe the event at S′, and we have the option whether we infer that it occurred at A′ or at B, *i. e.* (allowing for the FitzGerald contraction) at A; this is the place at which the testimony of a stationary observer would concur.

We will draw the diagram simply as fig. 7. One ought first of all to think of the definite hyperbola $x^2-y^2=\text{const}=\text{OA}^2$. This gives O and x_0. Then for a given u, one can get x' by drawing AM at the angle $\alpha=\sin^{-1}\frac{u}{c}$, and then,

since $x'^2-y'^2=\text{OA}^2$,

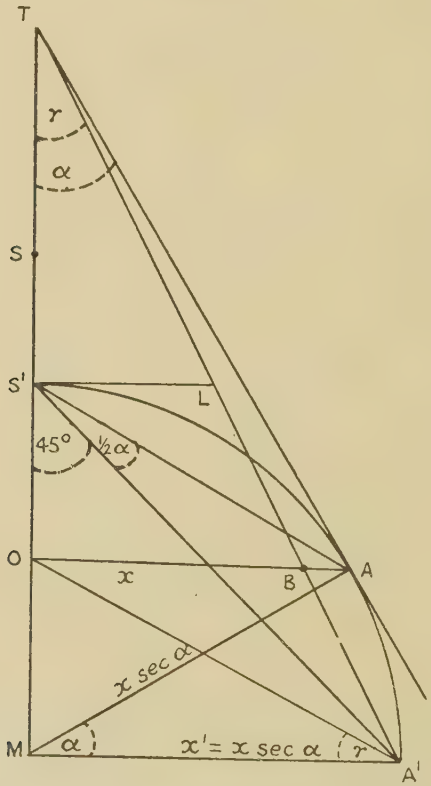
it follows that $x'^2 = OA^2 + OM^2$;

$= MA^2$;

$\therefore \quad MA' = MA$;

so practically we use a circle centre M, radius MA, to find A′ the desired point on the hyperbola. Then a tangent at A′ will give B, and the time which elapses

Fig. 7.



before the event A is seen is $\frac{OB}{c+u}$. There are several interesting pieces of geometry connected with the figure :—

(1) The line AT ⊥ to MA is at angle α with MOT, so TA′ is at angle γ, since

$$\tan A'TM = \frac{MA'}{MT} = \frac{MA}{MT} = \sin ATM = \sin \alpha = u/c.$$

(2) $OA'M=\gamma$, since its tangent$=\dfrac{OM}{MA'}=\dfrac{OM}{MA}=\sin\alpha$.

(3) A'T, A'O are equally inclined to the 45° line A'S'. Therefore, S'L (in fig. 7)=S'O, the triangles S'LA', S'OA' being congruent .

Or, putting it another way, there is a displacement of Jupiter due to relative velocity u so that its sloping line touches the hyperbola $x^2-(ct)^2=$constant instead of just sloping from the point $(x, 0)$ and therefore crosses the "now" line OA at a forward point B such that OB=OA cos α (FitzGerald contraction).

Now the eclipse effect reaches the eye in the time $\dfrac{x'+ct'}{c}$ from its actual stationary place at A$(x, 0)$. This time is equal to $\dfrac{x\cos\alpha}{c+u}$: *therefore* we have the right to choose between two alternatives :

(1) The eclipse seems to occur earlier and further away (at A') and the velocity of approaching light appears as just c and no more : *or*

(2) The eclipse occurs at time $t=0$, but by reason of relative motion the apparent distance is contracted into $x\cos\alpha$, and the light approaches with velocity c, but meets the observer with *relative* velocity $c+u$; in the time $\dfrac{x\cos\alpha}{c+u.}$.

We have to choose between these two alternatives, and we are inclined to think the extended FitzGerald contraction the less violating to common sense.

Caution.

But the contraction is generally supposed to apply to matter, and to be due to its relative motion. What the significance of applying it to space not in motion may be, demands further consideration. There is no doubt, however, that the travel of an observer with velocity u towards the light ought to make him perceive that it is moving relatively to him with the velocity $c+u$, and the contraction seems to follow as a kind of effect analogous to aberration.

We must remember that for *successive events* on Jupiter, the observer's track is up the line OT, and there are

different hyperbolæ, each event A being at the vertex, and A′ a point on it such that Jupiter's time-path touches the hyperbola at A′.

Every A will be on the line TA. Every A′ and every B will be on the line TA′. Also AOA′ is always γ, so as A is known, A′ is drawable by means of γ, *i. e.*, the OA′ lines are all parallel.

The advantage of B over A′ is that it does not alter the time of an event, but only contracts OA. The advantage of A′ over B is that it gives us quickly the time interval $\frac{MS'}{c}$, by means of MS′=MA′, using the 45° line from A′ to S′, S′ being the sight moment.

Note.—B is no help to the geometry, it just waits patiently for someone to pay attention to it.

LVII. *Notices respecting New Books.*

Origine des Rayons Gamma. Structure fine du spectre magnetique des Rayons alpha. By Salomon Rosenblum. No. 4 of Exposés de Physique Théorique under the direction of M. Louis de Broglie. [35 pages quarto.] (Paris: Hermann et Cie. Price 12 fr.)

THIS little monograph contains a concise account of the experimental work on the fine structure of alpha rays, in particular the work with the Paris electromagnets; and these experimental results are linked with the energy level schemes suggested by Gamow to explain the emission of gamma radiation. It is admirably arranged and printed, and since this work is one of the remarkable successes of commonplace ideas boldly applied to nuclear problems it should have a wide appeal to all physicists, not only to the specialist.

L'Existence du Neutron. By Irène Curie et F. Joliot. No. 2 of Exposés de Physique Théorique under the direction of M. Louis de Broglie. [22 pages quarto.] (Paris: Hermann et Cie. Price 6 fr.)

The two authors have here collected and set out an account of the series of experiments with ionization chamber and expansion chamber, which led to the discovery that the artificially excited nuclear radiation can project other nuclei.

As an account of an extremely valuable research it is very desirable, though one could wish that the title were justified by a fuller discussion of the neutron hypothesis, as has already been given very clearly by Dr. Chadwick. The printing and reproduction of the photographs are quite satisfactory.

Exposé de physique théorique.—I. *Sur une forme plus restrictive des relations d'incertitude.* D'après MM. Landau et Peierls. Par Louis de Broglie. (Pp. 24. 6 fr.)

V. *Mécanique quantique et causalité.* D'après M. Fermi. Par André George. (Pp. 18. 6 fr.) (Nos. 31 and 38 of Actualités scientifiques et industrielles. Paris: Hermann, 1932.)

These two booklets are both concerned with investigations connected with the indeterminacy principle. In each case a *résumé* with commentary, is given of a recent paper. That of Fermi is a discussion of the condition for exact determination of future events in quantum mechanics, while that of Landau and Peierls is concerned with the manner in which exactness of determination, of momentum for example, is further restricted by relativity conditions—considerations which have important bearing on present difficulties of nuclear physics. The presentation is admirably lucid, and the format of the booklets is pleasing.

The purpose of the series, under the general editorship of Louis de Broglie, is to present accounts of recent pieces of work which are of outstanding interest for theoretical physics, and which are likely to stimulate further investigation. One of the aims is to ease the difficulty of multiplicity of periodicals and languages by discussing selected important foreign scientific papers in separate booklets of the series. It may, however, be suggested that the appearance of even only a " petit fascicule " for each foreign paper worth reading seems hardly likely to simplify the problem of multiferous multilingual scientific publications ; and that the series might be more widely serviceable if the range of the individual booklets were not so restricted as in the examples under review —excellent though they are within that chosen range.

État Actuel de la Théorie du Neutron. By Jean-Louis Destouches. No. 3 of Exposés de Physique Théorique directed by M. Louis de Broglie. [Paper covers. 66 pages quarto.] (Paris: Hermann et Cie. Price 18 fr.)

The discovery of a neutral atomic particle has brought to our realization very effectively how much atomic theory depends on the electrical fields of the atomic particles.

The nature of the neutron, and the way in which it is to fit in our present scheme of atomic behaviour, is a problem which has quickly aroused the theoretical physicist, and in this admirably printed and arranged monograph the author has given a *résumé* of the attempts to reconcile the neutral particle with the necessities of atomic structure. The principal points discussed are the actual constitution of the neutron which plays so great a part in the further theoretical development (as for example, whether it obeys the Bose or the Fermi statistics), the emission of neutrons from nuclei, the difference between the effects of photons and of neutrons, and the passage of neutrons through matter. The book leaves one with the feeling that there could not be, as yet, a coherent theory of the neutron, and since the tentative speculations of the theoretician do not have any direct appeal except to the specialist, the more average physicist may not find much satisfaction in this review. To those who wish to study the neutron closely it will be an admirable way of breaking the ground.

Applications of the Absolute Differential Calculus. By A. J. McConnell. (Blackie & Son, Ltd. Price 20*s*.)

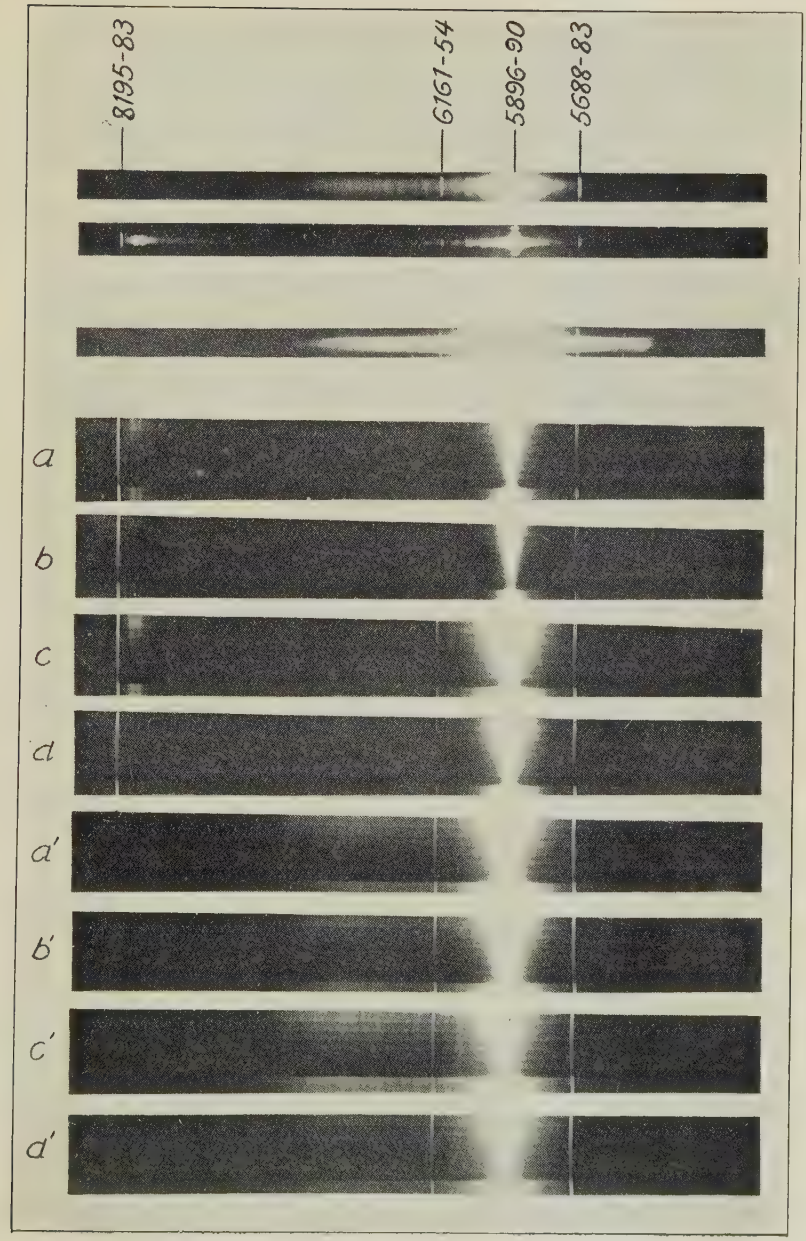
This book will satisfy a need which has long been felt by those interested in the Tensor Calculus. It is often assumed that this subject, owing to its use in the theory of Relativity, must be intrinsically difficult and must only have recondite applications. Professor McConnell has done good service in showing that Tensor Analysis is no more difficult to understand than classical Vector Analysis, which, indeed, it surpasses in the brevity of its notation and its economy in the use of symbols for its operations. Its applications, moreover, are much more extensive than is often imagined. More than one half of the book is devoted to algebraic and differential geometry treated by means of this calculus. The remainder deals with applied mathematics, and chapters are found on general dynamics, electromagnetism, elasticity, the special theory of relativity, etc. There is hardly a branch of applied mathematics, except perhaps the field of quantum mechanics, in which the Tensor Calculus cannot be applied with advantage. Students and teachers alike will be grateful to Professor McConnell for this guide to its uses, and the numerous exercises the book contains will undoubtedly add greatly to its value.

[*The Editors do not hold themselves responsible for the views expressed by their correspondents.*]

FIG. 1.

FIG. 2.

FIG. 3.



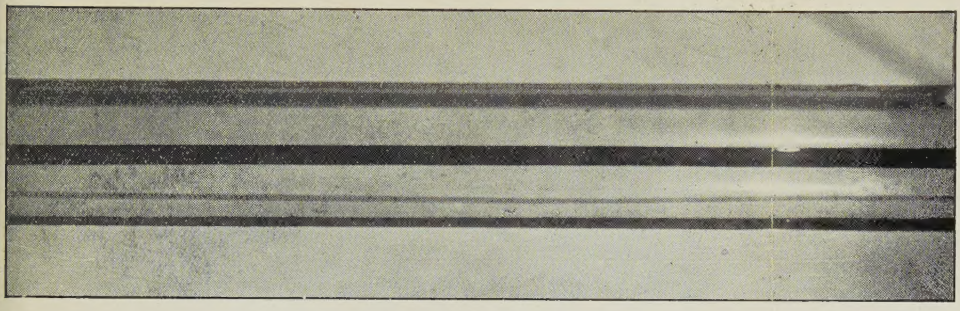


FIG. 11 (*a*).—Parallel streamline flow.

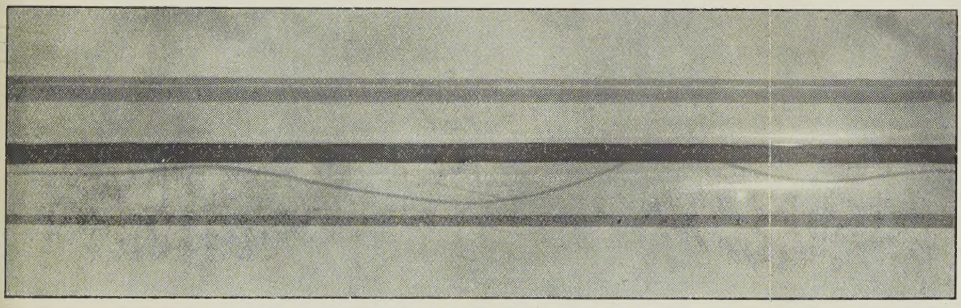


FIG. 11 (*b*).—The sinuous motion.

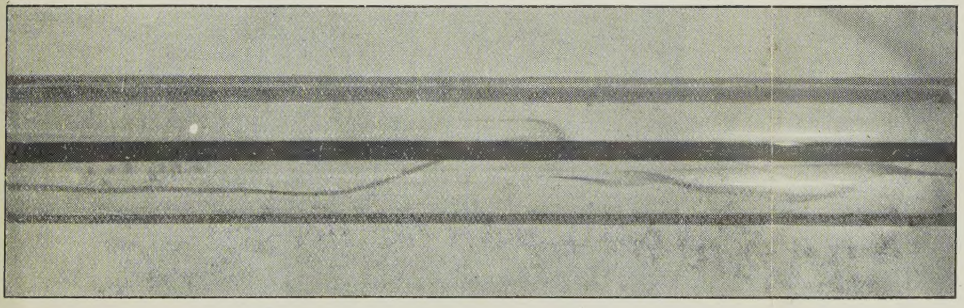


FIG. 11 (*c*).—Breaking up of sinuous motion.

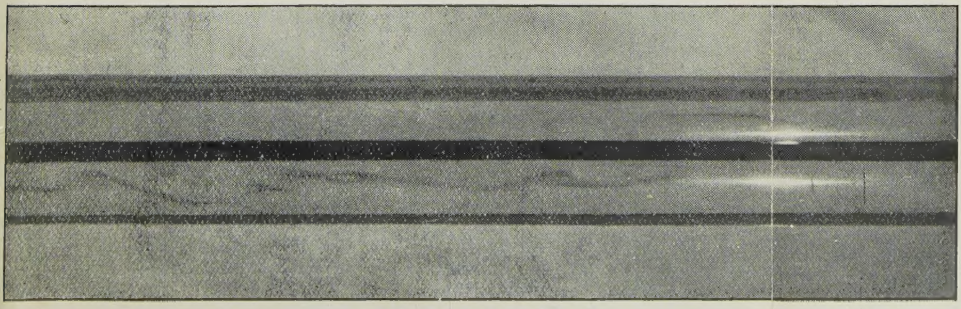
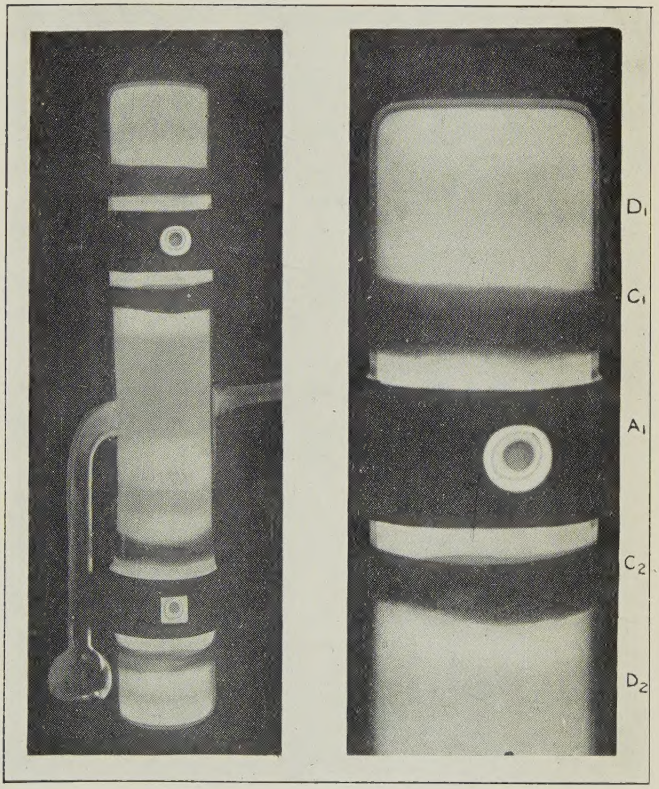


FIG. 11 (*d*).—Late stage, before general turbulence.

Fig. 1. Fig. 2.



Illustrating deposit of metallic mercury by high-frequency discharge.